

A

08/02/99
 3520 U.S. PRO

Certificate of Mailing

Date of Deposit August 2, 1999

Label Number: EL300789431US

I hereby certify under 37 C.F.R. § 1.10 that this correspondence is being deposited with the United States Postal Service as "Express Mail Post Office to Addressee" with sufficient postage on the date indicated above and is addressed to: BOX PATENT APPLICATION, Assistant Commissioner for Patents, Washington, D.C. 20231.

Sandra E. Marxen

Printed name of person mailing correspondence

Sandra E. Marxen
 Signature of person mailing correspondence

UTILITY PATENT APPLICATION TRANSMITTAL UNDER 37 CFR §1.53(b)

Attorney Docket Number	00786/246002
Applicant	David Moore <i>et al.</i>
Title	RETINOID X RECEPTOR-INTERACTING POLYPEPTIDES AND RELATED MOLECULES AND METHODS

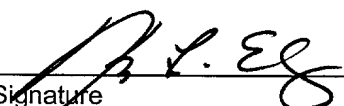
PRIORITY INFORMATION:

This application is a divisional of and claims priority from United States patent application 08/372,652, filed January 13, 1995.

APPLICATION ELEMENTS:

Cover sheet	1 page
Specification	43 pages
Claims	4 pages
Abstract	1 page
Drawing	20 pages
Combined Declaration and POA, which is: <input checked="" type="checkbox"/> Unsigned;	3 pages
<input checked="" type="checkbox"/> A copy from prior application 08/372,652 and the entire disclosure of the prior application is considered as being part of the disclosure of this new application and is hereby incorporated by reference therein.	2 pages
Statement Deleting Inventors	
Sequence Statement	
Sequence Listing on Paper	
Sequence Listing on Diskette	
Small Entity Statement, which is: <input checked="" type="checkbox"/> Unsigned	1 page
Preliminary Amendment	

3520 U.S. PRO
 09/365576
 08/02/99

IDS	
Form PTO 1449	
Cited References	
Recordation Form Cover Sheet and Assignment	
Assignee's Statement	
English Translation	
Certified Copy of Priority Document	
Return Receipt Postcard	1
FILING FEES:	
Basic Filing Fee: \$380	\$380.00
Excess Claims Fee: 27 - 20 x \$9	\$63.00
Excess Independent Claims Fee: 3 - 3 x \$39	\$0
Multiple Dependent Claims Fee: \$130	\$0
Total Fees:	\$443.00
<ul style="list-style-type: none"> ■ The filing fee is not being paid at this time. ■ Please apply any other charges, or any credits, to Deposit Account No. 03-2095. 	
CORRESPONDENCE ADDRESS:	
Karen L. Elbing, Ph.D. Reg. No. 35,238 Clark & Elbing LLP 176 Federal Street Boston, MA 02110	
Telephone: 617-428-0200 Facsimile: 617-428-7045	
 Signature	<u>2 August 1999</u> Date

Certificate of Mailing


Date of Deposit August 2, 1999

Label Number: EL300789431US

I hereby certify under 37 C.F.R. § 1.10 that this correspondence is being deposited with the United States Postal Service as **"Express Mail Post Office to Addressee"** with sufficient postage on the date indicated above and is addressed to BOX PATENT APPLICATION, Assistant Commissioner of Patents, Washington, D.C. 20231.

Sandra E. Marxen

Printed name of person mailing correspondence



Signature of person mailing correspondence

APPLICATION

FOR

UNITED STATES LETTERS PATENT

APPLICANT : DAVID MOORE, WONGI SEOL, AND HUENG-SIK CHOI

TITLE : RETINOID X RECEPTOR-INTERACTING POLYPEPTIDES AND
RELATED MOLECULES AND METHODS

Applicant or Patentee : David Moore *et al.*
Serial or Patent No. :
Filed or Issued : August 2, 1999
Title : RETINOID X RECEPTOR-INTERACTING POLYPEPTIDES AND RELATED MOLECULES AND METHODS

**VERIFIED STATEMENT (DECLARATION) CLAIMING SMALL ENTITY STATUS
 (37 CFR 1.9(f) and 1.27(d)) - NONPROFIT ORGANIZATION**

I hereby declare that I am an official empowered to act on behalf of the nonprofit organization identified below:

Name of Organization: Massachusetts General Hospital
Address of Organization: 55 Fruit Street, Boston, MA 02114
Type of Organization:

- ☐ University or Other Institution of Higher Education
☒ Tax Exempt under Internal Revenue Service Code (26 USC 501(a) and 501(c)(3))
☐ Nonprofit Scientific or Educational under Statute of State of the United States of America
 Name of State:
 Citation of Statute:
☐ Would Qualify as Tax Exempt under Internal Revenue Service Code (26 USC 501(a) and 501(c)(3)) If Located in the United States of America
☐ Would Qualify as Nonprofit Scientific or Educational under Statute of State of the United States of America If Located in the United States of America
 Name of State:
 Citation of Statute:

I hereby declare that the nonprofit organization identified above qualifies as a nonprofit organization as defined in 37 CFR 1.9(e) for purposes of paying reduced fees under section 41(a) and (b) of Title 35, United States Code with regard to the invention entitled RETINOID X RECEPTOR-INTERACTING POLYPEPTIDES AND RELATED MOLECULES AND METHODS by inventors David Moore, Wongi Seol, and Hueng-Sik Choi described in

- ☒ the specification filed herewith.
☐ application serial no. [**SERIAL NUMBER**], filed [**FILING DATE**].
☐ patent no. [**PATENT NUMBER**], issued [**ISSUE DATE**].

I hereby declare that rights under contract or law have been conveyed to and remain with the nonprofit organization with regard to the above identified invention.

If the rights held by the nonprofit organization are not exclusive, each individual, concern or organization having rights to the invention is listed below* and no rights to the invention are held by any person, other than the inventor, who could not qualify as a small business concern under 37 CFR 1.9(c) or by any concern which would not qualify as a small business concern under 37 CFR 1.9(d) or a nonprofit organization under 37 CFR 1.9(e).

*NOTE: Separate verified statements are required from each named person, concern or organization having rights to the invention averring to their status as small entities. (37 CFR 1.27)

Full Name:
Address:
☐ INDIVIDUAL ☐ SMALL BUSINESS CONCERN ☐ NONPROFIT ORGANIZATION

I acknowledge the duty to file, in this application or patent, notification of any change in status resulting in loss of entitlement to small entity status prior to paying, or at the time of paying, the earliest of the issue fee or any maintenance fee due after the date on which status as a small entity is no longer appropriate. (37 CFR 1.28(b))

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under section 1001 of Title 18 of the United States Code, and that such willful false statements may jeopardize the validity of the application, any patent issuing thereon, or any patent to which this verified statement is directed.

Name: David Glass, Ph.D.,

Title: Associate Director for Patents, Corporate Sponsored Research and Licensing

Address: Massachusetts General Hospital, Thirteenth Street, Building 149, Suite 1101, Charlestown, MA 02129

Signature: _____ **Date:** _____

RETINOID X RECEPTOR-INTERACTING POLYPEPTIDES
AND RELATED MOLECULES AND METHODS

Background of the Invention

5 This invention relates to receptor proteins.

This invention was made in part with Government funding, and the Government therefore has certain rights in the invention.

10 The retinoid X receptors (RXRs) are members of a large superfamily of intracellular hormone receptors. These proteins bind to specific DNA sequences and directly regulate transcription of target genes in response to activation by their specific ligands (Leid et al., Trends Biochem. Sci. 17:427-433, 1992; Leid et al., Cell 68:377-15 395, 1992; Mangelsdorf et al., Nature 345:224-229, 1990 and Yu et al., Cell 67:1251-1266, 1991). The RXRs belong to a large subgroup of the superfamily defined by a conserved subregion within the DNA binding domain. This group also includes the receptors for retinoic acid, thyroid hormone, 20 and vitamin D as well as a number of other less well characterized proteins, called orphan receptors, that do not have known ligands. As monomers, the members of this class can bind to sequences related to the hexameric consensus AGGTCA. RXR homodimers bind to tandem repeats of this 25 consensus separated by a single base pair (Mangelsdorf et al., Cell 66:555-561, 1991), and apparently to additional elements including β -RARE (Zhang et al., Nature 358:587-591, 1992). These homodimer binding sites confer specific response to 9-*cis*-RA (9-*cis*-RA), the ligand for the RXRs. 30 In addition, the RXRs heterodimerize with a variety of other family members, including the receptors for all- *trans*-retinoic acid, thyroid hormone (T3), and vitamin D. This

heterodimerization strongly increases the affinity of these receptors for their specific response elements (Yu et al., Cell 67:1251-1266, 1991; Zhang et al., Nature 358:587-591, 1992; Bugge et al., EMBO J. 11:1409-1418, 1992), and recent
5 evidence also demonstrates that it is also required for full hormone dependent transcriptional activity of at least the thyroid hormone receptor-RXR complex.

Mammals have three genes encoding α , β , and γ isoforms of RXR (Mangelsdorf et al., Genes & Dev. 6:329-344,
10 1992). The expression patterns of murine RXRs (Mangelsdorf et al., Genes & Dev. 6:329-344, 1992) and homologues of RXR found in *Xenopus* (Blumberg et al., Proc. Natl. Acad. Sci. USA 89:2321-2325, 1992) and *Drosophila* (Oro et al., Nature 347:298-301, 1990) suggest that the members of the RXR
15 family play important roles in several aspects of development and central nervous system differentiation as well as in adult physiology. Based on both their specific response to the 9-*cis*-RA metabolite and their heterodimerization with the RARs, it is clear that the RXRs
20 play a central role in the broad regulatory effects of retinoids. Moreover, their heterodimeric interactions with other family members indicate that the RXRs also play a central role in response to thyroid hormone, vitamin D, and perhaps other compounds. This dual function is unique
25 within the nuclear receptor superfamily.

Summary of the Invention

In a first aspect, the invention generally features a method for determining whether a test protein is capable of interacting with a retinoid X receptor (RXR) protein.

30 The method involves: (a) providing a host cell which contains (i) a reporter gene operably linked to a protein binding site; (ii) a first fusion gene which expresses a

first fusion protein, the first fusion protein including a retinoid X receptor protein covalently bonded to a binding moiety which is capable of specifically binding to the protein binding site; and (iii) a second fusion gene which expresses a second fusion protein, the second fusion protein including the test protein covalently bonded to a gene activating moiety; and (b) determining whether the test protein increases expression of the reporter gene as an indication of its ability to interact with the retinoid X receptor protein.

In a preferred embodiment, the method further involves treating the host cell with a ligand which binds the retinoid X receptor (preferably, 9-cis-RA) and identifying a ligand-dependent interacting protein by its ability to increase expression of the reporter gene upon treatment of the cell by the ligand. In another preferred embodiment, the method further involves treating the host cell with a ligand which binds the retinoid X receptor and identifying a ligand-independent interacting protein by its ability to increase expression of the reporter gene both in the presence and in the absence of ligand treatment. In yet another preferred embodiment, the method further involves treating the host cell with a ligand which binds the retinoid X receptor and identifying a ligand-sensitive interacting protein by its ability to increase expression of the reporter gene in the absence but not in the presence of ligand treatment.

In other preferred embodiments, the gene activating moiety is the gene activating moiety of B42.

In a second aspect, the invention features a substantially pure preparation of a retinoid X receptor (RXR)-interacting protein. Preferably, the RXR-interacting protein is RIP14, RIP15, RIP110, or RIP13; or includes an

amino acid sequence substantially identical to an amino acid sequence shown in any of Figs. 4, 5, 10, and 11 (SEQ ID NOS: 1-5); is derived from a mammal, for example, a human; binds a β -RARE site in the presence of RXR; or binds an EcRE site in the presence of RXR.

The invention also features purified DNA (for example, cDNA) which includes a sequence encoding an RXR-interacting protein, preferably encoding a human RXR-interacting protein (for example, the RXR-interacting proteins RIP14 (SEQ ID NO: 6, 14), RIP15 (SEQ ID NO: 7), RIP110 (SEQ ID NO: 8), or RIP13 (SEQ ID NO: 9)); a vector and a cell which includes a purified DNA of the invention; and a method of producing a recombinant RXR-interacting protein involving providing a cell transformed with DNA encoding an RXR-interacting protein positioned for expression in the cell; culturing the transformed cell under conditions for expressing the DNA; and isolating the recombinant RXR-interacting protein. The invention further features recombinant RXR-interacting protein produced by such expression of a purified DNA of the invention.

As used herein, "reporter gene" is meant a gene whose expression may be assayed; such genes include, without limitation, *lacZ*, amino acid biosynthetic genes, e.g. the yeast *LEU2* gene, luciferase, or the mammalian chloramphenicol transacetylase (CAT) gene. Reporter genes may be integrated into the chromosome or may be carried on autonomously replicating plasmids (e.g., yeast 2μ plasmids).

By "operably linked" is meant that a gene and a regulatory sequence(s) are connected in such a way as to permit gene expression when the appropriate molecules (e.g., transcriptional activator proteins or proteins which include transcriptional activation domains) are bound to the regulatory sequence(s).

By a "binding moiety" is meant a stretch of amino acids which is capable of directing specific polypeptide binding to a particular DNA sequence (i.e., a "protein binding site"). LexA represents a preferred DNA binding moiety in the invention. However, any other transcriptionally-inert or essentially transcriptionally-inert DNA binding domain may be substituted. The GAL4 DNA binding domain represents a somewhat less preferred DNA binding moiety for the system described herein.

By "gene activating moiety" is meant a stretch of amino acids which is capable of inducing the expression of a gene to whose control region it is bound. As used herein, a "weak gene activating moiety" is meant a stretch of amino acids which induces gene expression at a level below the level of activation effected by GAL4 activation region II (Ma and Ptashne, Cell 48:847, 1987) and is preferably at or below the level of activation effected by the B42 activation domain of Ma and Ptashne (Cell 51:113, 1987). Levels of activation may be measured using any downstream reporter gene system and comparing, in parallel assays, the level of expression stimulated by the GAL4- or B42-polypeptide with the level of expression stimulated by the polypeptide to be tested.

By "RXR-interacting protein" is meant a polypeptide which directly or indirectly physically interacts with a retinoid X receptor in the *in vivo* protein interaction assay described herein. Such an interaction may be hormone (or ligand) dependent or independent or may be hormone (or ligand) sensitive; it may also be transient in nature so long as the interaction is capable of producing a positive result in the interaction assay described herein. Preferably, such a polypeptide has an amino acid sequence which is at least 85%, preferably 90%, and most preferably

95% or even 99% identical to the amino acid sequence of an interacting protein described herein (e.g., RIP14, RIP15, RIP110, or RIP13) at the point of interaction with the retinoid X receptor, or at least 80% and preferably 90% identical overall.

By "substantially pure" is meant a preparation which is at least 60% by weight (dry weight) the compound of interest, i.e., an RXR-interacting protein. Preferably the preparation is at least 75%, more preferably at least 90%, and most preferably at least 99%, by weight the compound of interest. Purity can be measured by any appropriate method, e.g., column chromatography, polyacrylamide gel electrophoresis, or HPLC analysis.

By "purified DNA" is meant DNA that is not immediately contiguous with both of the coding sequences with which it is immediately contiguous (one on the 5' end and one on the 3' end) in the naturally occurring genome of the organism from which it is derived. The term therefore includes, for example, a recombinant DNA which is incorporated into a vector; into an autonomously replicating plasmid or virus; or into the genomic DNA of a prokaryote or eukaryote, or which exists as a separate molecule (e.g., a cDNA or a genomic DNA fragment produced by PCR or restriction endonuclease treatment) independent of other sequences. It also includes a recombinant DNA which is part of a hybrid gene encoding additional polypeptide sequence.

By "substantially identical" is meant an amino acid sequence which differs only by conservative amino acid substitutions, for example, substitution of one amino acid for another of the same class (e.g., valine for glycine, arginine for lysine, etc.) or by one or more non-conservative substitutions, deletions, or insertions located at positions of the amino acid sequence which do not destroy

the function of the protein (assayed, e.g., as described herein). Preferably, such a sequence is at least 85%, more preferably 90%, and most preferably 95% identical at the amino acid level to one of the sequences of Figs. 4, 5, 10, and 11 (SEQ ID NOS: 1-5). A "substantially identical" nucleic acid sequence codes for a substantially identical amino acid sequence as defined above.

By "transformed cell" is meant a cell into which (or into an ancestor of which) has been introduced, by means of recombinant DNA techniques, a DNA molecule encoding (as used herein) an RXR-interacting protein.

By "positioned for expression" is meant that the DNA molecule is positioned adjacent to a DNA sequence which directs transcription and translation of the sequence (i.e., facilitates the production of, e.g., an RXR-interacting protein).

By "purified antibody" is meant antibody which is at least 60%, by weight, free from the proteins and naturally-occurring organic molecules with which it is naturally associated. Preferably, the preparation is at least 75%, more preferably at least 90%, and most preferably at least 99%, by weight, antibody, e.g., RXR-interacting protein-specific antibody. A purified RXR-interacting protein antibody may be obtained, for example, by affinity chromatography using recombinantly-produced RXR-interacting protein and standard techniques.

By "specifically binds" is meant an antibody which recognizes and binds RXR-interacting protein but which does not substantially recognize and bind other molecules in a sample, e.g., a biological sample, which naturally includes RXR-interacting protein.

Other features and advantages of the invention will be apparent from the following detailed description thereof, and from the claims.

Description of the Drawings

5 The drawings are first described.

Figure 1 is a graph showing β -galactosidase expression conferred to LexA-RXR by B42-RIP clones. A yeast strain containing a *lacZ* reporter gene under the control of LexA binding sites and LexA-RXR was transformed with the indicated B42-fusion protein expression vectors. Strains coexpressing LexA-RXR and the indicated B42 fusions were grown overnight in liquid in the presence or absence of 10^{-6} M 9-*cis*-RA, added at the time of inoculation, and β -galactosidase activity was assayed.

15 **Figure 2** is a Northern analysis of RIP14 (panel A) and RIP15 (panel B). A Northern blot containing mRNAs from the indicated tissues (Clontech, Palo Alto, CA) was hybridized with RIP14 and RIP15 probes. In a longer exposure of RIP14, several larger bands not evident in this exposure were observed in liver and kidney.

20 **Figure 3** is a schematic representation of the structure of the RIP14 cDNA and protein isoforms. The original yeast clone isolated from the interaction trap is diagrammed at the top. The positions of the putative initiation and termination codons are indicated, and the 12 base pair (bp) addition in the D domain of RIP14-2 is shown as ‡. Various sequences specific to each isoform and a 63 bp repeat present in Clone 12, but not Clone 3 are indicated.

30 **Figure 4** is the nucleotide (SEQ ID NO: 6, 14) and deduced amino acid sequences (SEQ ID NOS: 1, 2) of the RIP14 clone. Numbers of nucleotides and amino acids of each

sequence are shown on the left side. For RIP14-1 (SEQ ID NO: 1), the sequence of Clone 6 is shown. Four amino acids unique to RIP14-2 (SEQ ID NO: 2) are also shown. The DNA binding domain (C domain) and a poly A signal are underlined. N-terminus unique to isoform 1 is shown as italicized letters and the putative initiation codons as bold letters. Because the 5' ends of the RIP14-2 clones are different, both Clone 3 and Clone 12 are shown. A 63 base pair direct repeat which occurs in the 5' end of Clone 12 is indicated as underlined and in italicized letters. The GenBank submission numbers of RIP14-1 is U09416. The submission number of RIP14-2 Clones 3 and 12 are U09417 and U09418, respectively.

Figure 5 is the nucleotide (SEQ ID NO: 7) and deduced amino acid sequences (SEQ ID NO: 3) of RIP15. The in frame termination codon in front of the initiation codon, the DNA binding domain (C domain), and a poly A signal are underlined. The GenBank submission number of this sequence is U09419.

Figure 6 is a photograph showing an SDS-polyacrylamide gel analysis of *in vitro* translated RIP proteins. Lane 1: RIP14-1, 2: RIP14-2 (No. 3), 3: RIP14-2 (No. 12), 4: RIP15, 5: positive control (luciferase, MW ca 69 Kd). Clone No. 15 was used for expression of RIP14-1, and the full length derivatives of both Clones No. 3 and 12 (as described herein) for RIP14-2. Molecular weight markers are shown.

Figure 7 is a schematic diagram showing a comparison of RIP14 and RIP15 sequences to other receptors. Percent sequence identities of RIP14 (panel A) and RIP15 (panel B) compared to DNA and ligand binding domains of the indicated receptor superfamily members are shown. For comparison, the isoform of each receptor member showing the highest score in

a GenBank database search was used. TR: mouse thyroid hormone receptor α -1 (accession number: P16416); RAR: zebrafish retinoic acid receptor γ (accession number: LO3400); RXR: mouse retinoic acid X receptor α (accession number: P28700); EcR: *Drosophila* ecdysone receptor (accession number: P34021). Lengths of the various domains of the receptors are not to scale.

Figure 8 is a series of photographs showing DNA binding by the RXR-interacting proteins. In panel A, RIP14-1 without (lane 2) or with RXR (lanes 3-5) was incubated with the hsp27 promoter (EcRE) end-labeled with [32 P]ATP. In panel B, RIP15 (lanes 2-5) and RIP14-1 (lanes 9-12) without (lanes 2 and 9) or with (lanes 3-5 and 10-12) RXR were incubated with β RARE end-labeled with [32 P]ATP. In both cases, a 50-fold molar excess of unlabeled specific (sp; EcRE, lane A4 and β RARE, lanes B4 and B11) or nonspecific (ns; AP1, lanes A5, B5, and B12) oligomers were added with the labeled probe. Incubations with probe alone are shown in lanes A1, B1, and B8. Cell lysates used for *in vitro* translation were also incubated with the probes without (lanes A6 and B6) or with (lanes A7 and B7) RXR. RIP14-2 and the RIP14-1 chimeras containing a four amino acid insertion in the D domain (RIP14C) were also incubated with β RARE and RXR as shown in lanes B13 and B14, respectively. Specific bands are indicated by arrows.

Figure 9 is a graph showing activity of RIP14-1, RIP14-2, and RIP-15 in mammalian cells. Vectors expressing intact RIPS, RXR, or the CDM vector were cotransfected into HepG2 cells as indicated herein with a luciferase reporter plasmid containing three copies of the β -RARE and pTKGH as an internal control. 9-*cis*-RA was added at 10^{-6} M approximately 18 hours after transfection. Results are luciferase expression normalized to the hGH internal

control. Consistent results were obtained in three independent experiments.

Figure 10 is the nucleotide (SEQ ID NO: 8) and deduced amino acid sequences (SEQ ID NO: 4) of RIP110.

5 Figure 11 is the nucleotide (SEQ ID NO: 9) and deduced amino acid sequences (SEQ ID NO: 5) of RIP13.

Detailed Description

Applicants have used an in vivo interaction trap system to identify and isolate proteins that physically
10 interact with retinoid X receptors and, in particular, with the ligand binding domain of the human RXR α . These proteins are termed RXR-Interacting Proteins (or RIPS). The isolation and characterization of exemplary RIPS now follows.

15 Isolation of proteins that interact specifically with RXR

Recently, several genetic methods have been used to identify and characterize protein-protein interactions (e.g., Fields et al., Nature 340:245-246, 1989; Gyuris et al., Cell 75:791-803, 1993). The principal idea of these
20 systems is that transcription activation and DNA binding are quite distinct functions within most eukaryotic transcription activators, generally localized to two separate domains. Many functional examples of chimeric transcriptional activators consisting of the DNA binding
25 domain of one protein attached to a heterologous activation domain have been characterized (Green et al., Nature 325:75-78, 1987; Ma et al., Cell 51:113-119, 1987). The fact that this attachment can be indirect, mediated by protein-protein interaction rather than the covalent linkage of separate
30 domains of a single protein forms the basis of the selection. A version of this type of system, called the interaction trap, has allowed the isolation of several new proteins that interact with several different targets,

including Max (Zervos et al., Cell 72:223-232, 1993), Cdc2 (Gyuris et al., Cell 75:791-803, 1993), and RAG-1 (Coumo et al., Proc. Natl. Acad. Sci. USA in press., 1994).

We used the interaction trap (Gyuris et al., Cell 5 75:791-803, 1993) to identify cDNAs encoding proteins that interact with the ligand binding domain of human RXR α . As shown in Figure 1, a chimeric protein consisting of the hinge (D) and ligand binding (E) domains of RXR α fused to the intact bacterial LexA repressor protein (LexA-RXR) is 10 not a strong transcriptional activator in yeast, in either the presence or absence of 9-*cis*-RA. However, LexA-RXR does activate expression from LexA binding sites in cells which also express a fusion protein consisting of a transcriptional activation domain joined to another protein, 15 such as the thyroid hormone receptor, which interacts specifically with RXR.

To isolate RXR-interacting proteins, we constructed a mouse liver cDNA library in a derivative of the yeast vector pJG4-5 (Gyuris et al., Cell 75:791-803, 1993; see 20 below), in which the cDNA sequences are fused to the B42 transcriptional activation domain (Ma et al., Cell 51:113-119, 1987). The liver was chosen because it is a major target organ for the actions of a number of nuclear hormone receptor superfamily members. This library was introduced, 25 as described herein, into a LexA-RXR expressing host in which transcription of both the β -galactosidase (β -gal) and LEU2 genes was under the control of LexA binding sites.

From 3×10^6 primary yeast transformants, a number of β -gal expressing colonies were identified in two independent 30 screens carried out in the presence or absence of 9-*cis*-RA. Since expression of the B42-cDNA fusion proteins was induced by growth on galactose, candidates obtained from each condition were tested for galactose-dependent expression of

both β -gal and LEU2, in the presence or absence of 9-*cis*-RA, using appropriate indicator plates (see below). To test the specificity of the interaction with RXR, cDNA plasmids were rescued from a number of candidates that showed appropriate galactose dependency and reintroduced into hosts expressing LexA alone or other LexA fusion proteins (e.g., LexA-Cdc2).

Candidates showing specific interaction with LexA-RXR were sequenced across the B42 fusion junction using an appropriate vector primer, and additional sequence was determined. The deduced amino acid sequences were compared to those in the GenBank and EMBL databases using the GCG (Devereux et al., Nucleic Acids Res. 12:387-395, 1984) and BLAST programs (Altschul et al., J. Mol. Biol. 215:403-410, 1990). This sequence comparison demonstrated that a number of the independently isolated clones encoded known RXR heterodimer partners, either PPAR (eight clones) or RAR α (six clones), providing strong confirmation of the specificity of the screening. All of these clones included intact ligand binding domains, as expected from the fact that the major heterodimerization function lies within that domain. For PPAR (Issemann et al., Nature 347:645-650, 1990), three clones started at amino acid 84, just N-terminal to the DNA binding (C) domain, four others started at 91, within the C domain, and one at 170, just past the C domain. For RAR (Leroy et al., EMBO J. 10:59-69), all six clones started at amino acid 132, within the C domain. Although previous results demonstrated that LexA-RXR interacts strongly with a chimera consisting of B42 fused to the thyroid hormone receptor, no B42-TR clones were obtained, presumably as a consequence of the very low level of expression of the TR mRNAs in liver.

Six clones representing three independent B42 fusions to the vitamin D binding protein (Yang et al.,

Genomics 7:509-516, 1990) were isolated. Since this secretory protein is unlikely to interact with the nuclear RXR protein in intact cells, it is difficult to explain why these clones were obtained. It is possible that the vitamin D binding protein shares some structural similarity with the vitamin D receptor, which is an RXR heterodimer partner, and that this structural conservation is the basis for the interaction. However, it is also possible that the interaction of this protein with RXR is simply an artifact of the yeast system, and these clones have not been studied further.

Several additional clones encoded novel proteins. Two, RIP14 and RIP15, were previously undescribed orphan members of the nuclear receptor superfamily. As with the PPAR and RAR isolates, the B42 fusion junctions in both cases were near the beginning of the hinge (D) domain that separates the DNA (C) and ligand binding (E) domains. Two other clones, RIP13 and RIP110, showed no significant similarity to any known protein and are candidate transcriptional co-activators.

The level of β -gal expression conferred by several of the RXR interactors was examined more quantitatively (as described herein). The results of β -galactosidase assays of liquid cultures of a series of appropriate strains are shown in Figure 1. As expected from previous results with B42-TR and numerous biochemical studies (Leid et al., Cell 68:377-395, 1992; Zhang et al., Nature 358:587-591, 1992), the interaction of B42-RAR with LexA-RXR was independent of the presence or absence of 9-*cis*-RA. Cells coexpressing LexA-RXR and the RIP14 and RIP15 chimeras showed levels of β -gal expression comparable to B42-RAR regardless of the presence or absence of the RXR ligand 9-*cis*-RA, indicating a relatively strong, ligand independent interaction with RXR.

With RIP13, the significant β -gal expression observed in the absence of 9-*cis*-RA was increased approximately three fold in the presence of the ligand. For RIP110, only basal levels of expression were observed in the absence of 9-*cis*-RA, but the level of β -gal expression was strongly induced in the presence of 9-*cis*-RA, indicating that the interaction of this protein with RXR is dependent on the presence of ligand.

Human cDNAs encoding any of these RIP polypeptides may be isolated using human cDNA libraries (for example, a human liver cDNA library) and standard techniques of hybridization.

Interaction of RIPs with other Receptors

Interactions of the RIP clones with other superfamily members, including RAR, TR, the glucocorticoid receptor (GR), and MB67, an orphan isolated in this lab (Baes et al., Mol. Cell. Biol. 14:1544-1552, 1994) were also tested using a series of appropriate LexA fusions. In particular, the experiments shown in Table 1 were carried out as follows. Yeast transformants containing a *lacZ* reporter gene under the control of LexA binding sites and the indicated B42- and Lex-fusion protein expression vectors were transferred onto galactose-Ura⁻His⁻Trp⁻ plates containing X-gal and incubated for two days. Relative levels of β -galactosidase activity were estimated and are depicted as follows: B, blue (strong interaction); LB, light blue (weak interaction); W, white (no interaction); nt, not tested. At least three separate colonies were tested on the same plate for interaction. To test the effects of ligands, 100 μ l of a 10⁻⁶M solution of the appropriate ligand was spread onto plates just before inoculation of cells (9-*cis*-RA for RXR, T3 for TR, and all-*trans*-RA for RAR). For B42-PPAR, the plasmid isolated from

the screening was used. Both full length and truncated RAR fusions to LexA were tested with identical results for all the B42 fusions except B42-110, which was tested only with the full length fusion.

As indicated in Table 1, RIP13 and RIP110 interacted with all proteins except glucocorticoid receptor (GR), whereas RIP14 and RIP15 interacted with RXR only.

Table 1

Lex Fusion	RXR		TR		RAR		MB67	GR	
ligand	+	-	+	-	+	-			
B42 fusion									
13	B	B	B	B	B	B	B		W
14	B	B	W	W	W	W	W		W
15	B	B	W	W	W	W	W		W
110	B	LB	B	W	B	B	B		W
PPAR	B	B	W	W	nt	nt	nt		nt
RXR	B	B	B	B	B	B	B		W

RIP13's interaction with all of the conventional receptors remained unaffected by the presence or absence of ligands. As with LexA-RXR, RIP110 interacted with LexA-TR in a ligand-dependent manner. Its interaction with LexA-RAR was not dependent on retinoic acid, however, and it also interacted constitutively with MB67, which does not have known ligands. This orphan showed substantial transcriptional activity in mammalian cells grown in the absence of any exogenously added ligands. The unique character of RIP13 and RIP110 in these interactions suggests that these proteins may have very important roles in

conserved functions of the nuclear hormone receptor
superfamily. In particular, the interaction of RIP13 with a
wide array of superfamily members indicates that it may be
involved in any of several processes that are common to a
5 number of different receptors. In addition to direct
transcriptional regulation, this could include nuclear
transport, for example. The ligand-dependent nature of the
interaction of RIP110 suggests that it may be directly
involved in ligand-dependent transcription or other
10 activities.

Isolation of Full-Length RIP14 and RIP15 cDNAs

Northern blot analysis as described herein revealed
that RIP14 is expressed only in liver and kidney in mouse,
predominantly as a broad band of approximately 1.8 to 2.2 kb
15 (Figure 2). Much lesser amounts of three or four higher
molecular weight species were also observed. In contrast,
an approximately 2.3 kb RIP15 mRNA was universally expressed
in a number of tissues. To obtain full-length clones of
these mRNAs, a mouse liver cDNA library was constructed and
20 screened by conventional hybridization with RIP14 and RIP15
probes. Eight separate clones were obtained for RIP14, and
four for RIP15.

All eight RIP14 clones were analyzed by digestion
with multiple restriction enzymes and either partial or
25 complete sequencing. As diagrammed in Figure 3, these
clones may be divided into two subgroups that apparently
encode distinct isoforms, referred to as RIP14-1 and RIP14-
2. Based on initiation at the first methionine of the open
reading frame, the RIP14-1 isoform is a protein of 484 amino
30 acids. Since there are no in frame termination codons
upstream from this methionine, it remains possible that this
isoform includes additional N-terminal sequence. The RIP14-
1 mRNA, however, is approximately 1.8 to 2.0 kb, as

demonstrated by the hybridization of an oligonucleotide probe specific for the 5' end of the RIP14-1 sequence to only the lower portion of the broad band recognized by a common probe, and the cloned sequences account for approximately 2 kb, assuming a 200 nucleotide poly A tail. Thus, if such an N-terminal extension exists, it must be minimal. As many members of the receptor superfamily have 5' untranslated regions of several hundred nucleotides that include additional upstream AUG codons, it is also possible that the correct RIP14-1 sequence begins with a methionine downstream of that indicated in Figure 4. For simplicity, we assume that the indicated reading frame is full length. As described below, results from in vitro translation are consistent with this assignment.

The RIP14-2 group differed from RIP14-1 in two respects. First, Clones 3 and 12 shared related 5' sequences that diverged from those of RIP14-1 at a position within the apparent RIP14-1 coding region. Most of the 5' sequence specific to clone 12 consisted of an additional copy of a 63 base pair sequence from the region common to the RIP14-2 class (Figures 3 and 4). The basis for this variation is not known. As expected, hybridization with a probe specific for these more extensive RIP14-2 sequences detected the upper portion of the broad band recognized by the common probe, indicating that the full length RIP14-2 mRNA is approximately 2.0 to 2.2. kb. This corresponds well to the approximately 2.2 kb predicted from the sequence obtained, plus a poly A tail. Because of the divergent sequence, the first methionine of the open reading frame of the 451 amino acid RIP14-2 isoform corresponds to residue 38 of the predicted RIP14-1 sequence. RIP14-2 also differed from RIP14-1 by an insertion of four amino acids located four amino acids downstream of the C-terminus of the DNA

binding domain. Although this insertion is only observed in clones that include the RIP14-2 type 5' end, it is possible that it is also present in a subset of mRNAs of the RIP14-1 type, and that it may be absent in a subset of mRNAs of the RIP14-2 type. However, there is currently no evidence for the existence of these two potential additional isoforms.

In contrast to the complex structure of RIP14, all of the clones of RIP15 show identical sequence, except one that is 64 base pairs longer at the 5' end. All four clones contain the same open reading frame for the predicted 446 amino acid RIP15 protein, with an in-frame terminator nine bases upstream of the initiation codon (Figure 5).

In vitro translation of transcripts for the two RIP14 isoforms produced a slightly bigger protein for RIP14-1 than for RIP14-2, approximately 57 and 55 Kd (Figure 6). This corresponds well to the predicted sizes of approximately 55 and 52 Kd, and is consistent with the assignment of the initiator codons. The calculated molecular weight of RIP15 is approximately 48 Kd. However, the major *in vitro* translation product was approximately 60 Kd (Figure 6). The basis for this apparently aberrant migration is unclear.

In sum, therefore, the RIP15 gene apparently encodes a single product, while the RIP14 gene expresses at least two closely related isoforms. By analogy with several examples of alternate promoter utilization for other superfamily members, it seems likely that this mechanism accounts for the distinct 5' ends of the two major classes of RIP14 cDNAs. Because the cDNAs appear to be full length, the alternative possibility that they are derived by alternative splicing of a common transcript from a single promoter seems less likely. In contrast to this relatively common N-terminal variation, the variation of the RIP14

isoforms within the D domain is apparently unique within the receptor superfamily. Without information on the structure of the RIP14 gene, it is not possible to be certain how this variation is generated. However, since the first 6

5 nucleotides of the inserted sequence match the consensus 5' or donor splice site (GU(A/G)AGU) (SEQ ID NO: 10), it is quite likely that it is a consequence of alternative donor site utilization.

10 In Figure 7, the amino acid sequences of the two orphans were compared to those of several other members of the nuclear hormone receptor superfamily. In the DNA binding (C) domain, RIP14 was most closely related to insect ecdysone receptors, sharing 82% sequence identity with that of *Drosophila* (Koelle et al., Cell 67:59-77, 1991), for
15 example. Interestingly, RIP15 is the superfamily member next most closely related to RIP14 in this domain, with 67% identity. The DNA binding domains of RIP15 and the ecdysone receptors share 64% identity, and these three sequences form a rather divergent subgroup within the superfamily. A
20 distinctive feature of the RIP15 C domain sequence is an insertion of two amino acids in the short region between the two zinc modules. Although they do not show particularly strong overall sequence identity with the RIP14/RIP15/ecdysone receptor subgroup elsewhere in the C
25 domain, a similar insertion is present in the thyroid hormone receptors.

Both RIP14 and RIP15 included matches to all of the conserved sequence motifs present in the putative ligand binding and dimerization (E) domains of other orphans and
30 conventional receptors (Seagraves et al., Genes & Dev. 4:204-219, 1990; Amero et al., Mol. Endocrinol 6:3-8, 1992; Laudet et al., EMBO J 11:1003-1013, 1992), including a conserved C-terminal sequence associated with ligand

dependent transcriptional activation (Danielian et al., EMBO J. 11:1025-1033, 1992). As with the C domain, overall comparisons based on ligand/dimerization domains placed both RIP14 and RIP15 in a divergent subgroup that also includes the ecdysone receptors. Within this region, RIP14 shares 42% and 37% identity with RIP15 and the ecdysone receptor, respectively, while RIP15 shares 42% identity with the ecdysone receptor. Overall, these three proteins are approximately as closely related to each other as the TRs are to the RARs.

Heterodimers of RIP14 and RIP15 with RXR Bind DNA
Specifically

To identify DNA sequences recognized by each orphan, *in vitro* translated proteins were used for gel shift assays. Because of their strong sequence similarity to the ecdysone receptor, a well studied ecdysone response element (EcRE) from the *Drosophila* hsp27 promoter (Riddilough et al., EMBO J. 6:3729-3734, 1987) was tested for binding to both orphans, in the presence and absence of RXR. This element consists of two hexamers that match the receptor binding consensus AGGTCA, arranged as an inverted repeat separated by one base pair (IR-1). As demonstrated in Figure 8, panel A, RIP14-1 bound to the EcRE, but only in the presence of RXR. The binding of the RIP14-2 isoform to this element was weaker than that of RIP14-1 when similar amounts of RIP14 proteins were used. RIP15 did not bind to the EcRE regardless of the presence or absence of RXR. Several other DNA elements were tested in the gel shift assay, including the retinoic acid response element from the promoter of the human RAR β 2 isoform (β RARE) (de The et al., Nature 343:177-180, 1990). In the presence of RXR, the β RARE was bound by both the RIP14 isoforms and RIP15 (Figure 8, panel B). Again, the binding of the RIP14-2/RXR heterodimer was weaker

than the RIP14-1/RXR heterodimer when a similar amount of RIP14 proteins were used (Figure 8, panel B, lanes 10 and 14). In contrast to the results with the EcRE, RIP14-1 showed some binding to the β RARE in the absence of RXR.

5 The difference in apparent binding affinity between RIP14-1 and RIP14-2 could be due to either the variation in the short A/B domains or the four additional amino acids in the D domain of RIP14-2. The former would be consistent with the recently reported effects of different A/B domains
10 on DNA binding by isoforms of the orphan ROR (Giguere et al., Genes & Dev. 8:538-553, 1994). The latter would be consistent with the fact that the insertion occurs within a region called the T box (Wilson et al., Science 254:107-
15 110, 1992), which is associated with effects on DNA binding to direct repeats by heterodimers of other superfamily members. A chimeric receptor containing the four additional amino acids (MYTG) of RIP14-2 at the corresponding region of the D domain of RIP14-1 was constructed and tested for binding to both the β RARE and EcRE in the presence of RXR.
20 As shown in Figure 8, panel B, lane 13, the binding of this chimeric protein (RIP14C) to the β RARE was similar to that of isoform 1 rather than isoform 2. The EcRE also showed a similar result. The lack of an effect of the insertion on RIP14-1 binding suggests that the differences in the A/B
25 domain may determine relative binding affinity of the two isoforms.

From the above results, we conclude that both RIP14 and RIP15 bind to an overlapping set of specific elements as heterodimers with RXR.

30 The at least partially overlapping DNA binding specificity of RIP14 and RIP15 is consistent with the similarity of their DNA binding domains and suggests that they may have overlapping functional roles. Since both

interact with the β RARE, it is possible that these functions include effects on the complex response to retinoids.

However, the inactivity of both intact and chimeric versions of the two orphans in transient transfections indicates that both require activation by binding of as yet unidentified ligands or by other processes.

Function of RIP14 and RIP15 *in vivo*

To test the transcriptional activity of RIP14-1, RIP14-2, and RIP15, vectors expressing each were cotransfected into HepG2 cells with a luciferase reporter plasmid in which three copies of the β RARE were inserted upstream of the TK promoter (Baes et al., Mol. Cell. Biol. 14:1544-1552, 1994; described herein). Expression from this reporter was transactivated more than 100 fold by RAR in the presence of retinoic acid, and 20-50 fold by the apparently constitutive orphan MB67 (Baes et al., Mol. Cell. Biol. 14:1544-1552, 1994). Neither of the two RIP14 isoforms or RIP15 was able to transactivate the β RARE reporter under any condition tested (Figure 9). This was confirmed using chimeras in which the A/B and DNA binding (C) domains of the thyroid hormone receptor (TR) were fused to the hinge (D) and ligand binding (E) domain of each orphan. In cotransfections of these chimeras with an analogous reporter plasmid containing two copies of the synthetic palindromic T3 response element (TREpal) (Brent et al., Mol. Endocrinol. 3:1996-2004, 1989), the TR-RIP14 chimera was not significantly different from the CDM vector alone. A two to three fold activation was observed with the TR-RIP15 chimera under a variety of conditions. However, this effect was quite modest by comparison to the more than 50 fold activation observed with TR in the presence of T3.

These transfection data suggest that both orphans may need specific ligands to activate transcription. A

number of compounds considered potential ligands for orphan receptors were tested, including several hydroxycholesterols, dehydroepiandrosterone (DHEA), α -tocopherol, thyroid hormone (T3), reverse T3, and several retinoids. No specific activity was observed with any of these compounds.

In cotransfections of the orphans with RXR α , basal expression was unaffected (Figure 9). In the presence of 9-*cis*-RA, cotransfection with RXR alone resulted in strong activation of the β RARE reporter (Figure 9). Previous results (Zhang et al., Nature 358:587-591, 1992) suggested that this effect was largely mediated by RXR homodimers, although heterodimers of RXR with endogenous RARs could also contribute. Cotransfection with RIP14-1 decreased 9-*cis*-RA induced expression by approximately 90%, and cotransfection with RIP15 blocked it completely (Figure 9). RIP14-2, which bound β RARE with somewhat lower apparent affinity, decreased the level of 9-*cis*-RA induced expression by 60% (Figure 9). These inhibitory effects could be due either to direct binding of inactive RIP/RXR heterodimers to β RARE, or to indirect effects of sequestration of RXR in complexes. In either case, the results confirmed that RIP proteins were expressed in such transfections and suggested that both orphans may be involved in the complex retinoid response.

RIP110 and RIP13

RIP110 and RIP13 cDNAs were sequenced by standard techniques and deduced amino acid sequences determined, also by standard techniques. These sequences are presented in Figures 10 and 11.

The materials and methods used in the above experiments are now described.

Strains and Plasmids

LexA fusion proteins were expressed from derivatives of the LexA fusion vector (LexA(1-202)+PL) (Gyuris et al., Cell 75:791-803, 1993), which expresses the intact LexA protein. The LexA-RXR and LexA-TR fusions included human
5 RXR α and rat TR β sequences extending from the C-terminal portion of the DNA binding domain to the C-terminus.

Analogous LexA fusions to RAR, MB67, and GR were constructed using PCR (polymerase chain reaction). For LexA-RAR, an additional fusion to the intact RAR α was also generated.

10 B42 fusion proteins were either isolated from the cDNA library as described below or inserted into a derivative of the vector pJG4-5 (Gyuris et al., Cell 75:791-803, 1993) using standard procedures. For *in vitro* translation, appropriate fragments were cloned into a previously
15 described bacteriophage T7 promoter expression vector (Carter et al., Mol. Cell. Biol. in press., 1994) and expressed using bacteriophage T7 RNA polymerase. Mammalian expression vectors were derivatives of CDM (Seed, Nature 329:840-842, 1987), and reporter plasmids were derivatives
20 of pTKluc (Carter et al., Mol. Cell. Biol. in press, 1994) in which the herpes virus TK promoter directs expression of luciferase.

Yeast strains were derivatives of EGY48 (MAT α leu2 trp1 ura3 his3 LEU2::pLexop6-LEU2 (Δ UAS LEU2)) (Gyuris et
25 al., Cell 75:791-803, 1993; Zervos et al., Cell 72:223-232, 1993), in which expression of the chromosomal LEU2 gene is under the control of LexA operators. EGY48 was successively transformed with 8H18-34 (Gyuris et al., Cell 75:791-803, 1993), in which expression of the *E. coli lacZ* (β -
30 galactosidase) gene is also under the control of *lexA* operators (selection for URA3⁺), and derivatives of the LexA fusion expression vector LexA(1-202)+PL (Gyuris et al., Cell 75:791-803, 1993) (selection for HIS3⁺).

cDNA Library Screening and Characterization of RXR
Interactors

An oligo(dT)-primed mouse liver cDNA library was constructed using standard procedures (Ausubel et al.,
5 Current Protocols in Molecular Biology, Greene Pub. Assoc. New York, 1994) in the plasmid cgatrp2 (selection for TRP1⁺), a derivative of the B42 expression vector pJG4-5. This plasmid also contains a tRNA suppressor gene, *supF*, which can be used to rescue only cgatrp2 (and not the other
10 two plasmids present in the yeast host) after transformation to an *Escherichia coli* strain containing a P3 plasmid. The library was amplified and used to transform an EGY48 derivative expressing LexA-RXR. 3x10⁶ primary yeast transformants were obtained from glucose-Ura⁻His⁻Trp⁻ plates and recovered as described (Gyuris et al., Cell 75:791-803,
15 1993). 2x10⁷ of these cells were plated on galactose-Ura⁻His⁻Trp⁻Leu⁻ plates with or without 9-*cis*-RA. About 100 LEU2 expressing colonies were selected for X-gal testing on galactose-Ura⁻His⁻Trp plates. Forty were chosen for further
20 analysis based on stable galactose dependency of both growth on Leu⁻ plates and expression of β -galactosidase (Gyuris et al., Cell 75:791-803, 1993; Zervos et al., Cell 72:223-232, 1993). The cDNA plasmids were recovered by transformation of *E. coli* MC1063/P3 and reintroduced into host strains
25 expressing LexA-RXR, LexA alone, or other chimeras such as LexA-Cdc2 (Gyuris et al., Cell 75:791-803, 1993) to test specificity of the interaction. Candidates that interacted specifically with LexA-RXR were selected and sequenced with a primer from the fusion site of the B42 transcription
30 domain by the standard dideoxynucleotide method. Based on sequence information and pattern of restriction endonuclease digestion, candidate clones were divided into several classes. In some cases, further sequence information was

obtained. The obtained sequences were used to search sequence databases. To isolate clones containing the full length RIP14 and RIP15 cDNAs, a mouse liver cDNA library constructed in the CDM8 plasmid by standard procedures was
5 screened by conventional hybridization methods with fragments of RIP14 and RIP15 which were [³²P]-labeled by random priming.

β-Galactosidase Assay of RXR-Interacting Clones

An EGY48 derivative containing the 8H18-34 *lacZ*
10 reporter plasmid was successively transformed with LexA and B42-fusion protein expression vectors to generate a series of strains coexpressing each LexA fusion with each B42 fusion. At least two separate colonies from glucose-Ura⁻ His⁻Trp⁻ plates were selected randomly for each coexpressing
15 strain and used to inoculate galactose-Ura⁻His⁻Trp⁻ liquid media to induce expression of the B42 fusion protein (Gyuris et al., Cell 75:791-803, 1993). Cultures were assayed for β-galactosidase as described (Ausubel et al., Current
20 Protocols in Molecular Biology, Greene Pub. Assoc., New York, 1994).

RNA Analysis

A Northern blot containing 2μg of polyA⁺ mRNA from the indicated tissues (Clontech, Inc., Palo Alto, CA) was hybridized with probes labeled by random priming using
25 standard procedures (Ausubel et al., Current Protocols in Molecular Biology Greene Pub. Assoc., New York, 1994).

Cell Culture and Transfections

HepG2 cells were propagated in Dulbecco's modified Eagle's medium containing 10% fetal bovine serum.
30 Transfections were carried out using calcium phosphate precipitation in the same medium, or in medium supplemented with charcoal stripped serum, as described (Baes et al., Mol. Cell. Biol. 14:1544-1552, 1994). HepG2 cells plated on

six well culture plates were cotransfected with 1 μ g of plasmids expressing intact RIPs, with or without 0.25 μ g of RXR α vector, and with 1.5 μ g of a reporter plasmid containing three copies of the β RARE (Baes et al., Mol. Cell. Biol. 14:1544-1552, 1994), and 2 μ g of pTKGH as an internal control. Luciferase activity was normalized using the level of growth hormone expressed from pTKGH. Each transfection was done in duplicate.

Proteins and Gel Shift Assay

RIP14 and RIP15 proteins were produced by *in vitro* translation (Promega TNT, Madison, WI) using expression vectors containing the RIP genes following a T7 promoter. To generate full length RIP14-2 constructs, fragments from RIP14-1 clone No. 15 containing regions from the middle of the E domain to the 3' end replaced the corresponding regions of RIP14-2 Clones No. 3 or 12, respectively. Sequences were confirmed by DNA sequencing. Human RXR α protein was expressed in *E. coli* using a bacterial expression vector based on the bacteriophage T7 promoter (Carter et al., Mol. Cell. Biol. in press., 1994). The oligonucleotides used for gel shift assays were as following: β -RARE, 5' gatccgggtagGGTTCaccgaaAGTTCActcga 3' (SEQ ID NO: 11); hsp27, 5'ctagacaagGGTTCAaTGCACTtggtccatcg 3' (SEQ ID NO: 12). Hexamers that match the AGGTCA (SEQ ID NO: 13) consensus half site or its complement are capitalized. Double stranded oligonucleotides were end-labeled using [³²P]ATP and kinase, and free nucleotide was removed by gel filtration. Proteins were preincubated with 20 μ l of gel shift assay buffer (10mM Tris (pH 8.0), 40mM KCl, 0.05% NP-40, 10% glycerol, 1mM DTT, 2.5mM MgCl₂ and 5ng of poly dI-dC) for 10 minutes in ice. This mixture was then combined with the indicated labeled probe and incubated for 20 minutes at room temperature. Specific or nonspecific

competitor oligomers were added with the probe. The mixtures were analyzed by 6% non-denaturing polyacrylamide gel electrophoresis using 0.5X Tris-Borate-EDTA (TBE) buffer at 4°C.

5 Identification of Ligands which Bind RXR-Interacting Proteins

Isolation of cDNAs encoding RXR-interacting proteins enables the identification and isolation of their ligands. Accordingly, one aspect of the invention features a
10 screening assay for the identification of compounds which specifically bind to the RXR-interacting proteins described herein. Such an assay may be carried out using a recombinant RXR-interacting protein.

In one example, the RXR-interacting protein
15 component is produced by a cell that naturally produces substantially none of the protein or by a cell which produces functionally deficient protein; suitable cells are, e.g., those discussed above with respect to the production of recombinant receptor, most preferably, mammalian cells
20 such as HepG2 cells. Host cells are transfected with (1) a vector which expresses a nucleic acid encoding the RXR-interacting protein (i.e., the "producer vector") and (2) a vector which includes an RXR-interacting protein binding site (e.g., for RIP14 and RIP15, the β RARE sequence
25 described herein) positioned upstream of a target gene which may be assayed (e.g., a CAT gene, a luciferase gene, or a β -galactosidase gene) (i.e., the "reporter vector"). Using a standard transactivation assay procedure (for example, the assay described herein), RXR-interacting protein activity is
30 assayed by measuring binding site-dependent target gene expression. Useful ligands are identified as those compounds which, when added to the host cell medium, effect a change in RXR-interacting protein-directed gene expression

(as detected using any reporter vector); useful ligands according to the invention may either increase or decrease RXR-interacting protein activity.

Any suitable transactivation technique, producer
5 vector, and binding site-containing reporter vector may be used. Descriptions of transactivation assays and generally useful vectors for the identification of ligands which bind other nuclear hormone receptors are described, e.g., in Evans et al. (U.S. Pat. No. 4,981,784, 1991); Evans et al.
10 (WO 90/07517); Evans et al. (WO90/01428); and WO88/03168; all hereby incorporated by reference. RXR-interacting proteins which may be used to screen for ligands include wild-type molecules as well as any appropriate chimeric protein, for example, those chimeric proteins described
15 herein.

Candidate ligands may be purified (or substantially purified) molecules or the ligand may be one component of a mixture of ligands (e.g., an extract or supernatant obtained from cells; Ausubel et al., supra). In a mixed ligand
20 assay, the RXR-interacting protein ligand is identified by testing progressively smaller subsets of the ligand pool (e.g., produced by standard purification techniques, e.g., HPLC or FPLC) until a single ligand is finally demonstrated to modulate the protein's activity. Candidate ligands
25 include peptide as well as non-peptide molecules.

Alternatively, a ligand may be identified by its ability to bind an RXR-interacting protein using affinity chromatography. Recombinant protein is purified by standard techniques, from cells engineered to express the protein
30 (e.g., those described above); the recombinant protein immobilized on a column (e.g., a Sepharose column or a streptavidin-agarose column by the immunoaffinity method of Ausubel et al., supra) and a solution containing one or more

candidate ligands is passed through the column. Such a solution (i.e., such a source of candidate ligands) may be, e.g., a cell extract, mammalian serum, or growth medium on which mammalian cells have been cultured and into which the cells have secreted factors (e.g., growth factors) during culture; again, candidate ligands include peptide as well as non-peptide molecules. A ligand specific for a recombinant RXR-interacting protein is immobilized on the column (because of its interaction with the protein). To isolate the ligand, the column is first washed to remove non-specifically bound molecules, and the ligand of interest is then released from the column and collected.

Ligands isolated by the above methods (or any other appropriate method) may, if desired, be further purified (e.g., by high performance liquid chromatography; see above). Once isolated in sufficiently-purified form, a novel peptide ligand may be partially sequenced (by standard amino acid sequencing techniques). From this partial amino acid sequence, a partial nucleic acid sequence is deduced which allows the preparation of primers for PCR cloning of the ligand gene (e.g., by the method of Ausubel et al., supra).

Identification of RXR-Interacting Protein DNA Binding Sites

Identification of the RXR-interacting proteins facilitates identification of their DNA binding sites. According to one approach, DNA binding sites may be identified using a gel shift assay, e.g., as described above for the identification of the RIP14 and RIP15 binding sites. Alternatively, a transactivation assay may be utilized. Briefly, candidate DNA binding sites are inserted upstream of a target gene whose expression may be assayed and the ability of an RXR-interacting protein to bind the DNA site

is assayed as its ability to activate downstream gene expression.

Alternatively, a DNA binding site may be identified by selectively retaining a protein-bound DNA fragment on a nitrocellulose filter. This approach relies on the ability of nitrocellulose to bind proteins but not double-stranded DNA. Purified RXR-interacting protein (e.g., purified by standard techniques from cells engineered to express the protein, e.g., those described above) is mixed with labelled double-stranded DNA (e.g., a random pool of DNA fragments) under conditions which allow interaction. After incubation, the mixture is suction-filtered through nitrocellulose, allowing unbound DNA to pass through the filter while retaining the protein and any DNA specifically bound to it. Bound DNA fragments are then eluted from the filter and analyzed by gel electrophoresis or amplification and cloning. A detailed description of this technique is published in Ausubel et al. (supra).

Candidate DNA fragments for either approach may be derived, for example, from a randomly cleaved or sonicated genomic DNA library, a randomly generated set of oligonucleotides, and/or may be derived from known nuclear hormone response elements (see, e.g., Evans et al., WO90/11273).

Identification of RXR-interacting protein DNA binding sites facilitates a search for the presence of such sites upstream of known or yet unidentified genes (e.g., by an examination of sequences upstream of known genes or by standard hybridization screening of a genomic library with binding site probes). RXR-interacting protein-mediated transcriptional control of genes bearing the binding site upstream may then be investigated (e.g., by transactivation

experiments as described above), potentially leading to the elucidation of novel RXR-interacting protein functions.

Chimeric Receptors

The functional domains of the RXR-interacting proteins may be swapped with the domains of other members of the nuclear hormone receptor family (see, e.g., Evans et al., WO 90/11273; Evans, *Science* 240:889, 1988) in order to produce receptors having novel properties. For example, fusion of an RXR-interacting protein DNA binding domain to the ligand-binding and gene activation domains of glucocorticoid receptor would confer hormonal regulation on genes downstream of RIP binding sites. Alternatively, fusion of an RXR-interacting protein DNA binding domain to a trans-repressing domain (see, e.g., Evans et al., WO90/14356) would result in repression of the basal level of expression of genes bearing upstream RIP binding sites. Examples of receptor domains which may be included in a chimeric RIP receptor are described in Evans et al. (WO 90/15815) and in Evans et al. (*Science* 240:889, 1988). Construction of receptor fusion genes is carried out by standard techniques of molecular biology.

Dominant Negative Mutants

Mutants of RXR-interacting proteins may be generated which interfere with normal RIP activity. Such mutants are termed "dominant negative" and fall into at least two classes: (a) ones which bind to their DNA binding site (thereby interfering with the ability of wild-type RXR-interacting protein to bind the same site) but which do not activate ligand-dependent gene expression and (b) ones which heterodimerize with other receptors (e.g., RXR) but which do not promote the biological response associated with the wild-type heterodimer.

The first class of RIP dominant negative mutants include those receptor polypeptides which contain a wild-type DNA binding domain and a mutant gene activation domain. Such mutants are unable to transactivate a reporter gene even in the presence of ligand (e.g., as measured using a CAT reporter gene with an upstream β RARE and the standard methods described above) but retain the ability to bind a RIP DNA binding site (as evidenced, e.g., by DNA footprint analysis using a β RARE DNA sequence; Ausubel et al., supra).

The second class of RIP dominant negative mutants include those receptor polypeptides which contain a wild-type heterodimerization domain. Such a mutant interacts with its heterodimer partner and disrupts the partner's function. In one particular example, a dominant negative RIP-interacting protein may be overproduced (e.g., by directing its expression from a very strong promoter); the abundant protein forms heterodimers with cellular RXR protein, soaking up available RXR and thereby preventing RXR homodimer formation as well as RXR heterodimer formation with other partner proteins (e.g., RAR, VDR, and T3R). Wild-type RXR-interacting protein may function as a dominant negative mutant if overproduced in this manner. However, a mutant RXR-interacting protein lacking gene activation function and/or a DNA binding domain is preferred.

Any of the above mutants may be generated by any method of random or site-directed DNA mutagenesis (see, e.g., Ausubel et al., supra).

Identification of Molecules that Modulate RXR-Interacting Protein Receptor Expression

Isolation of genes encoding RXR-interacting proteins also facilitates the identification of molecules which increase or decrease RIP expression. According to one approach, candidate molecules (e.g., peptide or non-peptide

molecules found, e.g., in a cell extract, mammalian serum, or growth medium on which mammalian cells have been cultured) are added at varying concentrations to the culture medium of cells which express RIP mRNA. RIP expression is then measured by standard Northern blot analysis (Ausubel et al., supra) using RIP cDNA as a hybridization probe. The level of RIP expression in the presence of the candidate molecule is compared to the level measured for the same cells in the same culture medium but in the absence of the candidate molecule. A molecule which promotes an increase or decrease in RIP expression is considered useful in the invention.

RXR-Interacting Protein Expression

In general, RXR-interacting proteins according to the invention may be produced by transformation of a suitable host cell with all or part of an RXR-interacting protein-encoding cDNA fragment (e.g., the cDNA described above) in a suitable expression vehicle.

Those skilled in the field of molecular biology will understand that any of a wide variety of expression systems may be used to provide the recombinant protein. The precise host cell used is not critical to the invention. The RXR-interacting protein may be produced in a prokaryotic host (e.g., *E. coli*) or in a eukaryotic host (e.g., *Saccharomyces cerevisiae* or mammalian cells, e.g., COS 1, NIH 3T3, or HeLa cells). Such cells are available from a wide range of sources (e.g., the American Type Culture Collection, Rockland, MD; also, see, e.g., Ausubel et al., *Current Protocols in Molecular Biology*, John Wiley & Sons, New York, 1994). The method of transformation or transfection and the choice of expression vehicle will depend on the host system selected. Transformation and transfection methods are described, e.g., in Ausubel et al. (*Current Protocols in*

Molecular Biology, John Wiley & Sons, New York, 1994); expression vehicles may be chosen from those provided, e.g., in *Cloning Vectors: A Laboratory Manual* (P.H. Pouwels et al., 1985, Supp. 1987).

5 One preferred expression system is the mouse 3T3 fibroblast host cell transfected with a pMAMneo expression vector (Clontech, Palo Alto, CA). pMAMneo provides: an RSV-LTR enhancer linked to a dexamethasone-inducible MMTV-LTR promotor, an SV40 origin of replication which allows
10 replication in mammalian systems, a selectable neomycin gene, and SV40 splicing and polyadenylation sites. DNA encoding an RXR-interacting protein would be inserted into the pMAMneo vector in an orientation designed to allow expression. The recombinant RXR-interacting protein would
15 be isolated as described below. Other preferable host cells which may be used in conjunction with the pMAMneo expression vehicle include COS cells and CHO cells (ATCC Accession Nos. CRL 1650 and CCL 61, respectively).

Alternatively, an RXR-interacting protein is
20 produced by a stably-transfected mammalian cell line. A number of vectors suitable for stable transfection of mammalian cells are available to the public, e.g., see Pouwels et al. (supra); methods for constructing such cell lines are also publicly available, e.g., in Ausubel et al.
25 (supra). In one example, cDNA encoding the RXR-interacting protein is cloned into an expression vector which includes the dihydrofolate reductase (DHFR) gene. Integration of the plasmid and, therefore, the RXR-interacting protein-encoding gene into the host cell chromosome is selected for by
30 inclusion of 0.01-300 μ M methotrexate in the cell culture medium (as described in Ausubel et al., supra). This dominant selection can be accomplished in most cell types. Recombinant protein expression can be increased by DHFR-

mediated amplification of the transfected gene. Methods for selecting cell lines bearing gene amplifications are described in Ausubel et al. (supra); such methods generally involve extended culture in medium containing gradually increasing levels of methotrexate. DHFR-containing expression vectors commonly used for this purpose include pCVSEII-DHFR and pAdD26SV(A) (described in Ausubel et al., supra). Any of the host cells described above or, preferably, a DHFR-deficient CHO cell line (e.g., CHO DHFR⁻ cells, ATCC Accession No. CRL 9096) are among the host cells preferred for DHFR selection of a stably-transfected cell line or DHFR-mediated gene amplification.

Once the recombinant RXR-interacting protein is expressed, it is isolated, e.g., using affinity chromatography. In one example, an anti-RXR-interacting protein antibody (e.g., produced as described herein) may be attached to a column and used to isolate the RXR-interacting protein. Lysis and fractionation of RXR-interacting protein-harboring cells prior to affinity chromatography may be performed by standard methods (see, e.g., Ausubel et al., supra). Alternatively, an RXR-interacting protein fusion protein, for example, an RXR-interacting protein-maltose binding protein, an RXR-interacting protein- β -galactosidase, or an RXR-interacting protein-trpE fusion protein, may be constructed and used for RXR-interacting protein isolation (see, e.g., Ausubel et al., supra; New England Biolabs, Beverly, MA).

Once isolated, the recombinant protein can, if desired, be further purified, e.g., by high performance liquid chromatography (see, e.g., Fisher, *Laboratory Techniques In Biochemistry And Molecular Biology*, eds., Work and Burdon, Elsevier, 1980).

Polypeptides of the invention, particularly short RXR-interacting protein fragments, can also be produced by chemical synthesis (e.g., by the methods described in *Solid Phase Peptide Synthesis*, 2nd ed., 1984 The Pierce Chemical Co., Rockford, IL).

These general techniques of polypeptide expression and purification can also be used to produce and isolate useful RXR-interacting protein fragments or analogs (described herein).

10 Anti-RXR-Interacting Protein Antibodies

Human RXR-interacting proteins (or immunogenic fragments or analogues) may be used to raise antibodies useful in the invention; such polypeptides may be produced by recombinant or peptide synthetic techniques (see, e.g.,
15 *Solid Phase Peptide Synthesis*, supra; Ausubel et al., supra). The peptides may be coupled to a carrier protein, such as KLH as described in Ausubel et al., supra. The KLH-peptide is mixed with Freund's adjuvant and injected into guinea pigs, rats, or preferably rabbits. Antibodies may be
20 purified by peptide antigen affinity chromatography.

Monoclonal antibodies may also be prepared using the RXR-interacting proteins described above and standard hybridoma technology (see, e.g., Kohler et al., *Nature* 256:495, 1975; Kohler et al., *Eur. J. Immunol.* 6:511, 1976;
25 Kohler et al., *Eur. J. Immunol.* 6:292, 1976; Hammerling et al., *In Monoclonal Antibodies and T Cell Hybridomas*, Elsevier, NY, 1981; Ausubel et al., supra).

Once produced, polyclonal or monoclonal antibodies are tested for specific RXR-interacting protein recognition
30 by Western blot or immunoprecipitation analysis (by the methods described in Ausubel et al., supra). Antibodies which specifically recognize an RXR-interacting protein are considered to be useful in the invention; such antibodies

may be used, e.g., in an immunoassay to monitor the level of RXR-interacting protein produced by a mammal (for example, to determine the subcellular location of any of these retinoid X receptor interacting proteins.

5 Preferably, antibodies of the invention are produced using fragments of the RXR-interacting protein which lie outside highly conserved regions and appear likely to be antigenic, by criteria such as high frequency of charged residues. In one specific example, such fragments are
10 generated by standard techniques of PCR and cloned into the pGEX expression vector (Ausubel et al., supra). Fusion proteins are expressed in *E. coli* and purified using a glutathione agarose affinity matrix as described in Ausubel, et al., supra). *Current Protocols in Molecular Biology*
15 (Greene Pub. Assoc., New York, 1994)). To attempt to minimize the potential problems of low affinity or specificity of antisera, two or three such fusions are generated for each protein, and each fusion is injected into at least two rabbits. Antisera are raised by injections in
20 a series, preferably including at least three booster injections.

 Antisera is cleared of anti-GST antibodies using GST immobilized on a glutathione column, and the antisera are checked by ELISA for titer and specificity, using GST fusion
25 proteins as controls. Antisera is also checked for its ability to immunoprecipitate in vitro translated RXR-interacting proteins or control proteins, such as glucocorticoid receptor, CAT, or luciferase. Western blots of total or nuclear versus cytoplasmic fractionated HeLa
30 cell proteins are also probed with the antisera to assess specificity and to characterize subcellular compartmentalization. In these and other immunologic

assays, specificity is confirmed by the specific competition with the GST fusion protein.

Once the specificity of an antiserum is confirmed, it may be used in any standard indirect immunofluorescence procedure to determine the subcellular distribution of the RXR-interacting protein in a particular cell type.

Use

The proteins described herein interact with retinoid X receptor and are thus likely to mediate or modulate RXR function. In particular examples, RIP14 and RIP15 block RXR-dependent activation of β -RARE linked genes, and such proteins (or peptides derived from these proteins, particularly, short peptides which are capable of RXR interaction), may facilitate the production of pharmacologic modifiers of RXR function. Such therapeutic polypeptides of the invention may be administered by any appropriate route, e.g., intravenously, at a dosage which is effective to modulate RXR function. Treatment may be repeated as necessary for alleviation of disease symptoms.

The polypeptides of the invention are also useful for identifying those compartments of mammalian cells which contain proteins important to the function of the retinoid X receptor. Antibodies specific for a particular RXR-interacting protein may be produced as described above. The normal subcellular location of the protein is then determined either in situ or using fractionated cells by any standard immunological or immunohistochemical procedure (see, e.g., Ausubel et al., supra; Bancroft and Stevens, Theory and Practice of Histological Techniques, Churchill Livingstone, 1982).

Antibodies specific for RXR-interacting proteins also find diagnostic use in the detection or monitoring of

RXR-related diseases. Levels of an RXR-interacting protein in a sample may be assayed by any standard technique. For example, its expression may be monitored by standard Northern blot analysis or may be aided by PCR (see, e.g.,
5 Ausubel et al., supra; PCR Technology: Principles and Applications for DNA Amplification, ed., H.A. Ehrlich, Stokton Press, NY). These techniques are enabled by the provision of the RXR-interacting protein sequences described herein. Alternatively, standard immunological or
10 immunohistochemical procedures (e.g., those described above) may also be used with the antibodies described herein for RXR-interacting protein detection.

Other Embodiments

In other embodiments, the invention includes any
15 protein which is substantially identical to a human RXR-interacting protein (Figs. 4, 5, 10, and 11; SEQ ID NOS: 1-5); such homologs include other substantially pure naturally occurring mammalian RXR-interacting proteins (for example, human RXR-interacting proteins) as well as allelic variants;
20 natural mutants; induced mutants; proteins encoded by DNA that hybridizes to the RXR-interacting protein DNA sequence of any of Figs. 4, 5, 10, and 11 (SEQ ID NOS: 6-9, 14) under high stringency conditions or low stringency conditions (e.g., washing at 2X SSC at 40°C with a probe length of at
25 least 40 nucleotides); and polypeptides or proteins specifically bound by antisera directed to an RXR-interacting protein, especially by antisera to the RXR binding domain of the RXR-interacting protein. The term also includes chimeric polypeptides that include an RXR-
30 interacting protein fragment.

The invention further includes analogs of any naturally occurring RXR-interacting protein. Analogs can

differ from the naturally occurring RXR-interacting protein by amino acid sequence differences, by post-translational modifications, or by both. Analogs of the invention will generally exhibit at least 85%, more preferably 90%, and
5 most preferably 95% or even 99% identity with all or part of a naturally occurring RXR-interacting protein sequence. The length of comparison sequences will be at least 15 amino acid residues, preferably at least 25 amino acid residues, and more preferably more than 35 amino acid residues.

10 Modifications include *in vivo* and *in vitro* chemical derivatization of polypeptides, e.g., acetylation, carboxylation, phosphorylation, or glycosylation; such modifications may occur during polypeptide synthesis or processing or following treatment with isolated modifying
15 enzymes. Analogs can also differ from the naturally occurring RXR-interacting protein by alterations in primary sequence. These include genetic variants, both natural and induced (for example, resulting from random mutagenesis by irradiation or exposure to ethanemethylsulfate or by site-specific mutagenesis as described in Sambrook, Fritsch and
20 Maniatis, Molecular Cloning: A Laboratory Manual (2d ed.), CSH Press, 1989, or Ausubel et al., supra). Also included are cyclized peptides molecules and analogs which contain residues other than L-amino acids, e.g., D-amino acids or
25 non-naturally occurring or synthetic amino acids, e.g., β or γ amino acids.

In addition to full-length polypeptides, the invention also includes RXR-interacting protein fragments. As used herein, the term "fragment," means at least 20
30 contiguous amino acids, preferably at least 30 contiguous amino acids, more preferably at least 50 contiguous amino acids, and most preferably at least 60 to 80 or more contiguous amino acids. Fragments of RXR-interacting

proteins can be generated by methods known to those skilled in the art or may result from normal protein processing (e.g., removal of amino acids from the nascent polypeptide that are not required for biological activity or removal of
5 amino acids by alternative mRNA splicing or alternative protein processing events).

Preferable fragments or analogs according to the invention are those which facilitate interaction of the peptide with a retinoid X receptor.

10 All publications and patent applications mentioned in this specification are herein incorporated by reference to the same extent as if each independent publication or patent application was specifically and individually indicated to be incorporated by reference.

15 Other embodiments are within the following claims.
What is claimed is:

Claims

1 1. A method for determining whether a test protein
2 is capable of interacting with a retinoid X receptor (RXR)
3 protein, comprising:

4 (a) providing a host cell which contains

5 (i) a reporter gene operably linked to a
6 protein binding site;

7 (ii) a first fusion gene which expresses a
8 first fusion protein, said first fusion protein comprising a
9 retinoid X receptor protein covalently bonded to a binding
10 moiety which is capable of specifically binding to said
11 protein binding site; and

12 (iii) a second fusion gene which expresses a
13 second fusion protein, said second fusion protein comprising
14 said test protein covalently bonded to a gene activating
15 moiety; and

16 (b) determining whether said test protein increases
17 expression of said reporter gene as an indication of its
18 ability to interact with said retinoid X receptor protein.

1 2. The method of claim 1, wherein said method
2 further comprises treating said host cell with a ligand
3 which binds said retinoid X receptor and identifying a
4 ligand-dependent interacting protein by its ability to
5 increase expression of said reporter gene upon treatment of
6 said cell by said ligand.

1 3. The method of claim 1, wherein said method
2 further comprises treating said host cell with a ligand
3 which binds said retinoid X receptor and identifying a
4 ligand-independent interacting protein by its ability to

5 increase expression of said reporter gene both in the
6 presence and in the absence of said ligand treatment.

1 4. The method of claim 1, wherein said method
2 further comprises treating said host cell with a ligand
3 which binds said retinoid X receptor and identifying a
4 ligand-sensitive interacting protein by its ability to
5 increase expression of said reporter gene in the absence but
6 not in the presence of said ligand treatment.

1 5. The method of claim 1, wherein said gene
2 activating moiety is the gene activating moiety of B42.

1 6. The method of claim 2, wherein said ligand is 9-
2 cis-RA.

1 7. A substantially pure RXR-interacting protein.

1 8. The protein of claim 7, comprising an amino acid
2 sequence substantially identical to the amino acid sequence
3 of RIP14-1 shown in Figure 4 (SEQ ID NO: 1).

1 9. The protein of claim 7, comprising an amino acid
2 sequence substantially identical to the amino acid sequence
3 of RIP14-2 shown in Figure 4 (SEQ ID NO: 2).

1 10. The protein of claim 7, comprising an amino
2 acid sequence substantially identical to the amino acid
3 sequence of RIP15 shown in Figure 5 (SEQ ID NO: 3).

1 11. The protein of claim 7, comprising an amino
2 acid sequence substantially identical to the amino acid
3 sequence of RIP110 shown in Figure 10 (SEQ ID NO: 4).

1 12. The protein of claim 7, comprising an amino
2 acid sequence substantially identical to the amino acid
3 sequence of RIP13 shown in Figure 11 (SEQ ID NO: 5).

1 13. The protein of claim 7, wherein said
2 polypeptide is derived from a mammal.

1 14. The protein of claim 13, wherein said mammal is
2 a human.

1 15. The protein of claim 13, wherein said
2 polypeptide binds a β -RARE site in the presence of RXR.

1 16. The protein of claim 13, wherein said
2 polypeptide binds an ECRE site in the presence of RXR.

1 17. Purified DNA comprising a sequence encoding a
2 protein of claim 7.

1 18. The purified DNA of claim 17, wherein said DNA
2 encodes a human RXR-interacting protein.

1 19. The DNA of claim 17, comprising a DNA sequence
2 substantially identical to the DNA sequence of RIP14-1 shown
3 in Figure 4 (SEQ ID NO: 6).

1 20. The DNA of claim 17, comprising a DNA sequence
2 substantially identical to the DNA sequence of RIP14-2 shown
3 in Figure 4 (SEQ ID NO: 14).

1 21. The DNA of claim 17, comprising a DNA sequence
2 substantially identical to the DNA sequence of RIP15 shown
3 in Figure 5 (SEQ ID NO: 7).

1 22. The DNA of claim 17, comprising a DNA sequence
2 substantially identical to the DNA sequence of RIP110 shown
3 in Figure 10 (SEQ ID NO: 8).

1 23. The DNA of claim 17, comprising a DNA sequence
2 substantially identical to the DNA sequence of RIP13 shown
3 in Figure 11 (SEQ ID NO: 9).

1 24. A vector comprising the purified DNA of claim
2 17.

1 25. A cell containing the purified DNA of claim 17.

1 26. A method of producing a recombinant RXR-
2 interacting protein comprising,
3 providing a cell transformed with DNA encoding an
4 RXR-interacting protein positioned for expression in said
5 cell;
6 culturing said transformed cell under conditions for
7 expressing said DNA; and
8 isolating said recombinant RXR-interacting protein.

1 27. RXR-interacting protein produced by expression
2 of the purified DNA of claim 17.

RETINOID X RECEPTOR-INTERACTING POLYPEPTIDES
AND RELATED MOLECULES AND METHODS

Abstract of the Disclosure

Disclosed is a method for determining whether a test protein is capable of interacting with a retinoid X receptor protein. The method involves: (a) providing a host cell which contains (i) a reporter gene operably linked to a protein binding site; (ii) a first fusion gene which expresses a first fusion protein, the first fusion protein including a retinoid X receptor protein covalently bonded to a binding moiety which is capable of specifically binding to the protein binding site; and (iii) a second fusion gene which expresses a second fusion protein, the second fusion protein including the test protein covalently bonded to a gene activating moiety; and (b) determining whether the test protein increases expression of the reporter gene as an indication of its ability to interact with the retinoid X receptor protein. Also disclosed is purified DNA encoding retinoid X receptor-interacting proteins and the polypeptides expressed from such DNA.

92287.B11

FIG. 1

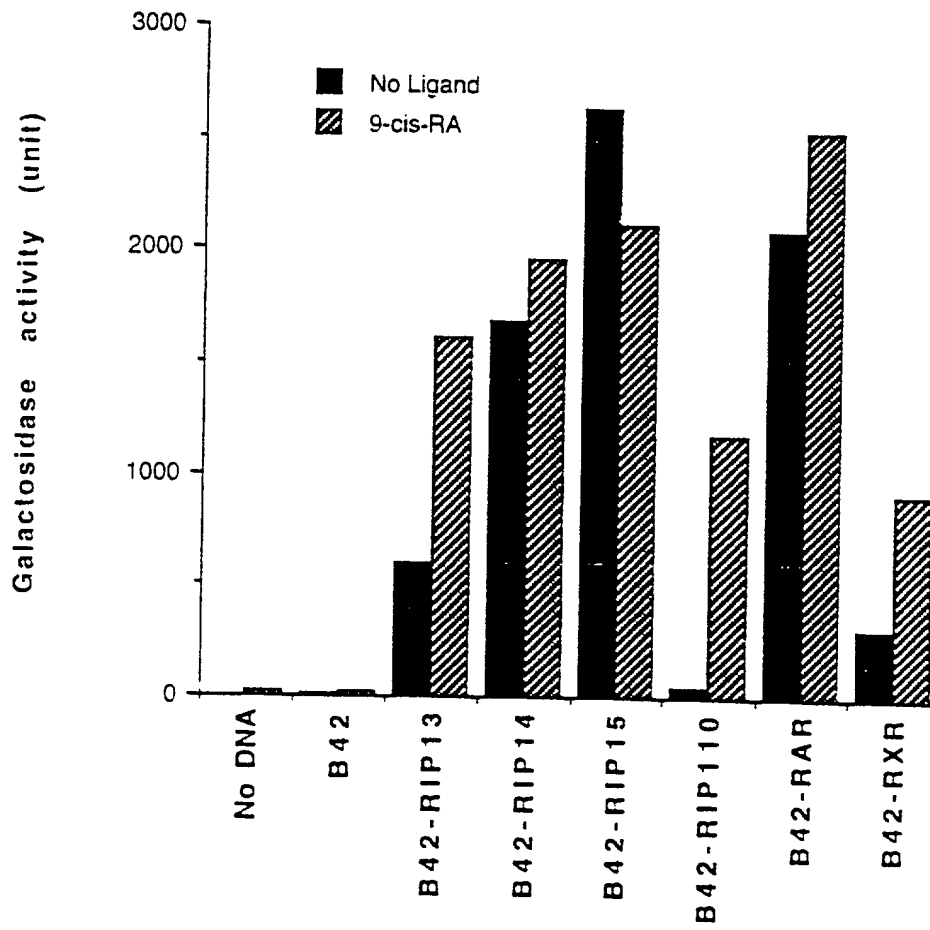


FIG. 2

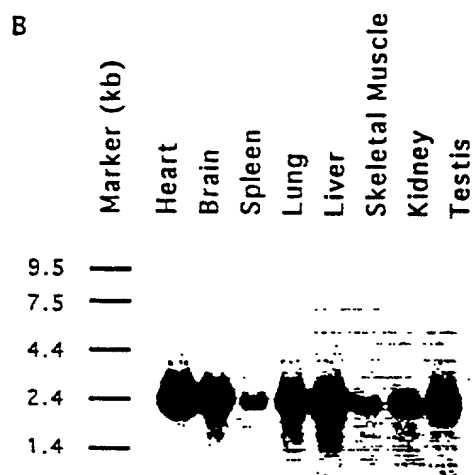
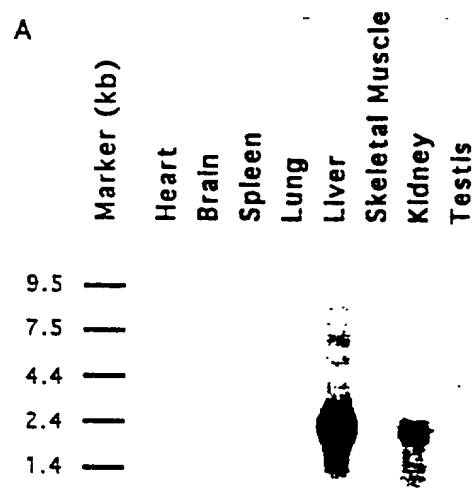


FIG. 3

Yeast clone

No. 6

No. 15

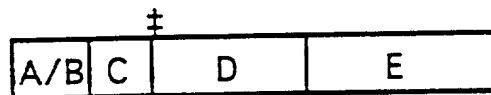
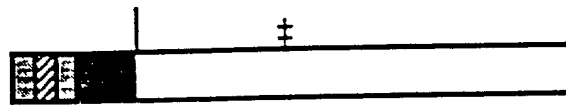
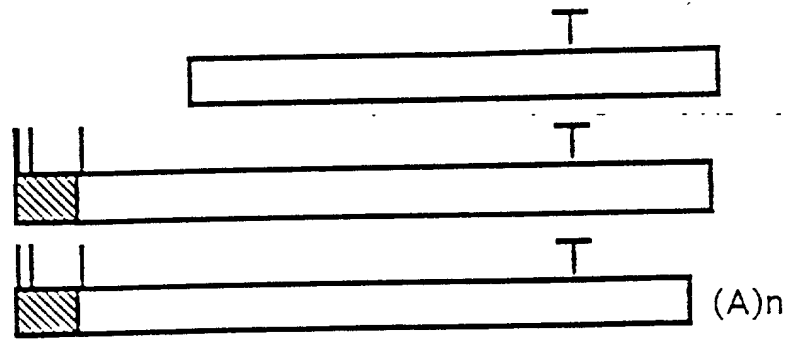
RIP14-1

No. 3

No. 12

No. 19

RIP14-2



- | ATG codon
- T Termination codon
- ‡ 12 bp addition
- ▨ 63 bp repeat

Specific to RIP14-2 clone 3

TGGTCA	CCCAGGCTTCGTGTTCA	GTTCTCTCTCCTCCTTCTCTCCTCCTCAGCCCCACTGTCTCTCT	60	-	-
CATCTCA	CCCC		71	-	-

Specific to RIP14-2 clone 12

AAAAATTAATGCGGGTACTAGAAAGGTAAGATCGGCGTTCGGTATCTCTCTTTAGAGCGGTGGACAT - 60
CTCACCCG - 68

Common to RIP14-2

AGAGAAGAACCGAGTTCGAGAGTCTACAGC	<u>AAAAATTA</u>	<u>CTGGGCACTAGAAAGG</u>	<u>AAGACT</u>	131	128	-
<u>GGGCTCCGAATCCTCTT</u>	<u>TAGAGCGCTTGGACAT</u>	<u>CTTGGCCCAAGCAATCCAAGGATCTTA</u>		191	188	-
<u>TFTGAGGACCACCATCCCAAGA</u>	<u>GTACTTCTCAAGGTTG</u>	<u>AAAGTTGGAGTGGTAGCCAA</u>		251	248	-
<u>GATGAATCTGATTGGGCAC</u>	<u>TCCATTTACAGGCTACGGACGAGTTTCTCTTTCTGAAAGC</u>			311	308	-
<u>TTATTT</u>				317	314	-

Specific to RIP14-1

GGAGCTAAGGATGGTGATGCAGTTTCAGGGCTTAGAAAATCCAATTCAGATTAGTCTTC	-	-	60
M V M Q F Q G L E N P I Q I S L	-	-	16
ACCACAGCCACCGGCTGTCAGGATTTGTGCCGGAACGGGATGAGTGTGAAGCCAGCTAAA	-	-	119
H H S H R L S G F V P D G M S V K P A K	-	-	36

Common to both isoforms

GGTATGCTAACAGAACACGCGGGCAGGGCCCTCTGCGGGCAGATCTCGGATTTGGGAATCGTACT	378	375	180
G M L T E H A A G P L G Q N L D L E S Y	19	19	56
CCCCATACAACAATGTCCCGTTTCTCAAGTTCAGCCACAGATTTCTCTCTCGTCTTACT	438	435	240
S P Y N N V P F P Q V Q P Q I S S S S Y	39	39	76
ATTCCACCGTGGGCTTCTACCCCCAACACCGGAGAGCTGGTATTCTCTCGGCATCTATG	498	495	300
Y S N L G F Y P Q Q P E D W Y S P G I Y	59	59	96
AACTCAGGGGATGCCCCGTGAGACTGGGTACCAGGGAGAGACTGAGGTATCAGAGATGC	558	555	360
E L R R M P A E T G Y Q G E T E V S E M	79	79	116
CTGTGACAAAGAGCGCGGATGGCGCGGGCATGGGCAGGCAGATAAAGCGGATGAGC	618	615	420
P V T K K P R M A A A S A G R I K G D E	99	99	136
TGTGTGTGTGTGTGTGGAGACAGGGGCTGTGGGTACCATACAAAGCGGTCACCTGTGAGG	678	675	480
L Q N V C G D F A S G Y H Y N A L T C E	119	119	156

GCTGCAAGGTTTCTTCCGAAGAAGCATTACCAAGAACGCCGTGTACAAGTGTAAAGAACG	738	735	540
G C K G F F R R S I T Y N A V Y K C K N	139	139	176
GGGGCAACTGGGTGATGGACATGTACATGGCGAGGAAGTGGCAGGAGTGGCCGGCTAAGGA	798	795	600
G G N C V M D M Y M R R K C O E C R L R	159	159	196
<div style="display: flex; align-items: center; justify-content: center;"> <div style="border: 1px solid black; padding: 5px; margin-right: 10px;"> GTATGTATACAG M Y T G </div> RIP14-2 specific </div>			
AGTGCAGAGAGATGGGGATGTTGGCTGAAT GTTTGTTAACTGHAATCCAGTGTAAATCTA	870	867	660
F C P E M G M L A E C L L T E I Q C K S	183	183	216
AACGGCTAAGGAAAATGTGAAGCAGCAGCGCTGATCAGACAGTGAATGAGGACGACAGCG	930	927	720
K R L R K N V K Q H A D Q T V N E D D S	203	203	236
AAGGGCGTGACTTGGGACAAGTGAACCTCCACAACCAAGTTTTGCAGGGAGAAAACGGAAC	990	987	780
E G R D L R Q V T S T T K F C R E K T E	223	223	256
TCACGGCAGACCAGCAGACCCTCCTGGATTATATTATGGATTCTGTACAACAAACAGAGAA	1050	1047	840
L T A D Q Q T L L D Y I M D S Y N K Q R	243	243	276
TGCTTCAGGAAATCACAATAAAATCTTAAAAGAAGAATTTAGTGCAGAAGAAAATTTTC	1110	1107	900
M P Q E I T N K I L K E E F S A E E N F	263	263	296
TCATATTAACAGLAATGGCAACCAGCCATGTACAGATTCTCGTAGAATTCACAAAAAGC	1170	1167	960
L I L T E M A T S H V Q I L V E F T K K	283	283	316
TTCCAGGGTTTCAGACACTGGACCACGAAGATCAGATTGCTTTTGCTCAAAGGGTCCGCAG	1230	1227	1020
L P G F Q T L D H E D Q I A L L K G S A	303	303	336
TGGAGGCCATGTTTCTTCTGTTCCGGGGAGATTTTCTAATAGAACTTCTGCGCGGACATG	1290	1287	1080
V E A M F L R S A E I F N K K L P A G H	323	323	356
CAGACCTGTTGGAAGAAAGAAATTCGAAAGAGTGGTATCTCTGATGAGTATATAACCCCGA	1350	1347	1140
A D L L E E R I R K S G I S D E Y I T P	343	343	376
TGTTTCAGTTTCTATAAAAGTGTGGAGAACTCAAAATGACTCAGGAGGAGTACGGCTCTGC	1410	1407	1200
M F S F Y K S V G E L K M T Q E E Y A L	363	363	396
TCACAGCGATCGTCATCCTCTCTCCAGACAGACAATACATCAAGGACAGAGAGGGCGGTGG	1470 (1453)	1260	
L T A I V I L S P D R Q Y I K D R E A V	383 (378)	416	
AGAAGCTGCAGGAGCCCCCTGCTTGATGTGCTACAAAAGCTGTGCAAGNTGTACCAGCCTG	1530	-	1320
E K L Q E P L L D V L Q K L C K M Y Q P	403	-	436
AGAACCACAGCATTTCGGCTGCTCTCGGGTCGGCTGACGGAACTCCGGACATTCAACC (1554)	-	-	1380
E N P Q H F A C L L G R L T E L R T F N (411)	-	-	456
ATCACCACGGTCAGATGCTGATGTCTTGGAGAGTGAATGATCACAAGTTCACCCCGCTCC	-	-	1440
H H H A E M L M S W R V N D H K F T P L	-	-	476
TCTGTGAGATCTGGGATGTGCASTGATGGACACCACTGGGGCTGGCTCCTTGTCTCTCTC	-	-	1500
L C E I W D V Q *	-	-	484

GGAACAGAAACCTTGTTTCGTTTGTACCTGGTTTCACTCAAGAATCTCAATGAATATTTA	-	-	1560
TGTGGCAATTATACACCTCCACGGTTGTAAATACAGACTAGATAGAAGTGGTTTCCCCA	-	-	1620
CACTGTATTTTACAAGGCTTCAGGAAACCCCACTGGCATGCCCTTTTGGCCTAATTAAAT	-	-	1680
CAATTGTTACTTCAATTCTATCTACTGAGCTAGGGGCATATTATTCTTCATTGACAATA	-	-	1740
TTATATATATTTTATAAAGTTGAGCTGTTTTCAACTGAGACAATAAA	-	-	1787

1560
1570
1580
1590
1600
1610
1620
1630
1640
1650
1660
1670
1680
1690
1700
1710
1720
1730
1740
1750
1760
1770
1780
1790
1800

GCACGGSCAACAGAGTGGGAGACCCCTGGCACCCCTGCCGATCGGTCAGTCAT -71
GAGCCCCGCTCCCCCTGGTGCAOGGAGAGGGGCGGGGCTTGAACAAGCAGGCTGCTTC -11
GTGACCCACTATGTCTTCCCCCACAAGTTCTCTGGACACTCCCGTGCTGGGAATGGTTC 50
M S S P T S S L D T P V P G N G S 17
TCTCAGCCCCAGTACCTCCGCCACGTCAOCCACTATTAAGGAAGAGGGGCAGGAGACTGA 110
P Q P S T S A T S P T I K E E G Q E T D 37
TCTCTCTCCAGGCTCTGAAGGGTCCAGCTCTGCTACATCGTGGTTCATCTTAGAGCCAGA 170
P P P G S E G S S S A Y I V V I L E P E 57
GGATGAGCTTGAGGGCAAGCGGAAGAAGGGGCGGGGCGGGAAGATGCTGGGCCATGAGCT 230
D E P E R K R K K G P A P K M L G H E L 77
GTGCGCGGTGTGCGGAGACAAGGCTTCGGGCTTCCACTACAACGTGCTCAGCTGTGAAGG 290
C R V C G D K A S G F H Y N V L S C E G 97
CTGCAAAGGCTTCTTCCGGCGCAGTGTGGTCCACGGTGGGGCGGGGCGCTATGCTGTGCG 350
C K G F F R R S V V H G G A G R Y A C R 117
GGGCAGCGGAACCTGOCAGATGGATGCTTCATGCGGCGCAAGTGCCAGCTCTGCGGGCT 410
G S G T C O M D A F M R R K C O L C R L 137
GCGCAAGTGCAAGGAGGCTGGCATGCGGGAGCAGTGGTGCTCTCTGAGGAGCAGATTGCG 470
R K C K E A G M R E Q C V L S E E Q I R 157
GHAGAAAAGGATTTCAGAAGCAGCAACAGCAGCAGCCACCCATCTGAGCCAGCAGC 530
K K R I Q K Q Q Q Q Q P P P P S E P A A 177
CAGCAGCTCAGGCGCGGCAGCGGCTCCCCCTGGCACTTCGGGAAGCAAGCAGCCAGGGCTC 590
S S S G R P A A S P G T S E A S S Q G S 197
CGGGGAAGSAGAGGSCATCCAGCTGAOCGCGGCTCAGGAGCTGATGATCCAGCAGTTAGT 650
G E G E G I Q L T A A Q E L M I Q Q L V 217
TGCCGCGCAGCTGCAGTGCAACAAACGATCTTTCTCCGACCAGCCAAAGTCACGCCCTG 710
A A Q L Q C N K R S F S D Q P K V T P W 237
GCCCCCTGGGTGCAGACCTCAGTCCCGAGATGCCCCGTACGCAACGCTTTGCCCACTTCAC 770
P L G A D P Q S R D A R Q Q R F A H F T 257
CGAGCTAGCCATCATCTCGGTCCAGGAGATTGTGGACTTTTGCCCAAGCAGGTGCCAGGGTT 830
E L A I I S V Q E I V D F A K Q V P G F 277
CTTGCAAGTTGGGCGGGGAGGAACAGATCGCCCTCTGAAGGCGTCCACCATTGAGATCAT 890
L Q L G R E D Q I A L L K A S T I E I M 297
GTTGCTAGAACAGCCAGAGCTACAAACCGAGACAGATGCATCAGTTCTTGAAGGA 950
L L E T A R R Y N H E T E C I T F L K D 317

CTTCACCTACAGCAAGGACGACTTCCACCGTGCAGGCTTGCAGGTGGAATTCATCAATCC 1010
F T Y S K D D F H R A G L Q V E F I N P 337

CATCTTCGAGTTCTCGCGGGCCATGCGGCGGCTGGGCCTGGACGATGCAGAGTATGCCCTT 1070
I F E F S R A M R R L G L D D A E Y A L 357

GCTTATCGCCATCAACATCTTCTCAGCCGATCGGCGCTAATGTGCAGGAGCCCAGCCGTGT 1130
L I A I N I F S A D R P N V Q E P S R V 377

GGAGGCCCTGCAGCAGCCCTACGTGGAGGCGCTCCTCTCCTACACGAGGATCAAGCGCCC 1190
E A L Q Q P Y V E A L L S Y T R I K R P 397

ACAGGACCAGCTCCGCTTCCACGGCATGCTCATGAAGCTGGTGAGCCTGGGCACCCCTCAG 1250
Q D Q L R F P R M L M K L V S L R T L S 417

CTCCGTGCACTCGGAGCAGGTCTTTGCAITGGGACTCCAGGACAAGAAGCTGCGGCCCTT 1310
S V H S E Q V F A L R L Q D K K L P P L 437

GCTGTCCGAGATCTGGGATGTGCACGAGTAGGGGCAGCCACAAGTGCCCCAGCCTTGGTG 1370
L S E I W D V H E * 446

GTGTCTTCTGAAGATGGACTCTTCACTCTCCTCCTGGGGTGGGAGGACATTGTACCGG 1430

CCCAGTCCCTCGGGCTCAGCCTCAAACCTCAGCGGCAGTTGGCACTAAGAAGGCCCCACCC 1490

CACCCATTGAGTCTTCCAAGAGTGGTGAGGGTCACAGGTCTTAGCCTCTGACCGTTCCCA 1550

GCTGCCCTCCCACCCACGCTTACACCTCAGCCTACCACACCATGCACCTTGAGTGGAGAG 1610

AGGTTAGGGCAGGTGGCCCCCCCCACAGTTGGGAGACCACAGGCCCTCTCTTCTGCCCTTT 1670

TATTTTAAATAAAAAACAATAAAGTTTGAGTACAAGCCAAAAAAAAAAAAAAAAAAAAA 1730

JUN 25 1993 12:53:50

FIG. 6

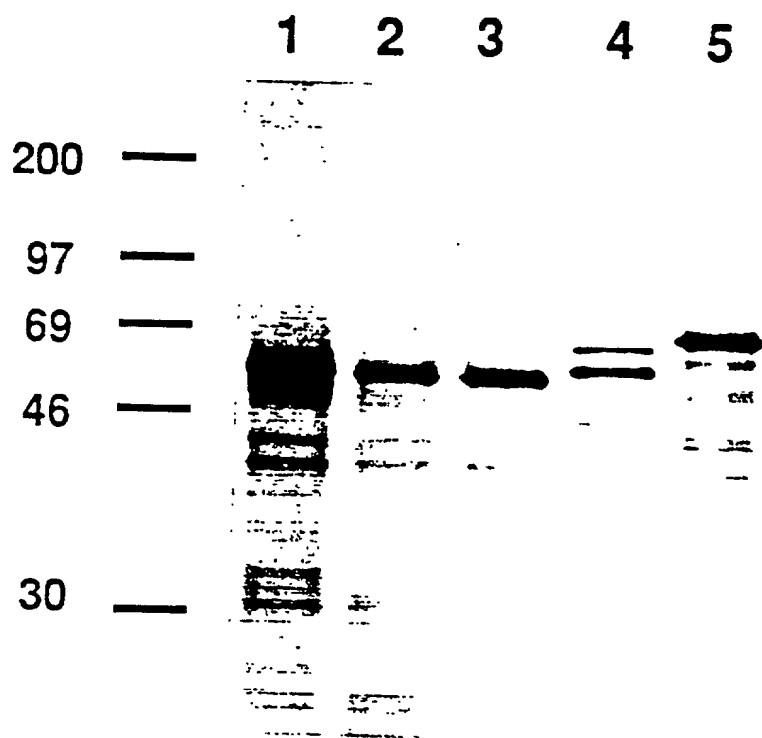


FIG. 7

A.

RIP14		DNA		LIGAND/DIMER
EcR		82		37
RIP15		67		42
TR		61		26
RAR		59		31

B.

RIP15		DNA		LIGAND/DIMER
EcR		64		42
RAR		56		38
RXR		51		27
TR		50		32

FIG. 8

		RIP14-1				cell lysate	
RXR	-	-	+	+	+	-	+
Competitor	-	-	-	sp	ns	-	-



1 2 3 4 5 6 7

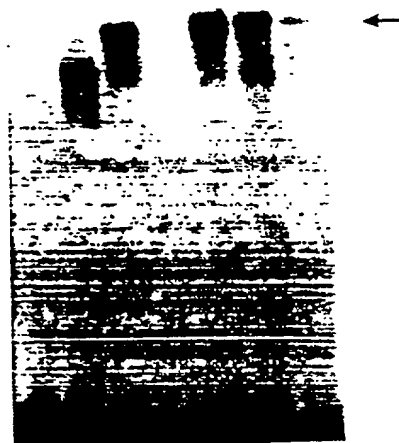
B

		RIP15				cell lysate	
RXR	-	-	+	+	+	-	+
Competitor	-	-	-	sp	ns	-	-



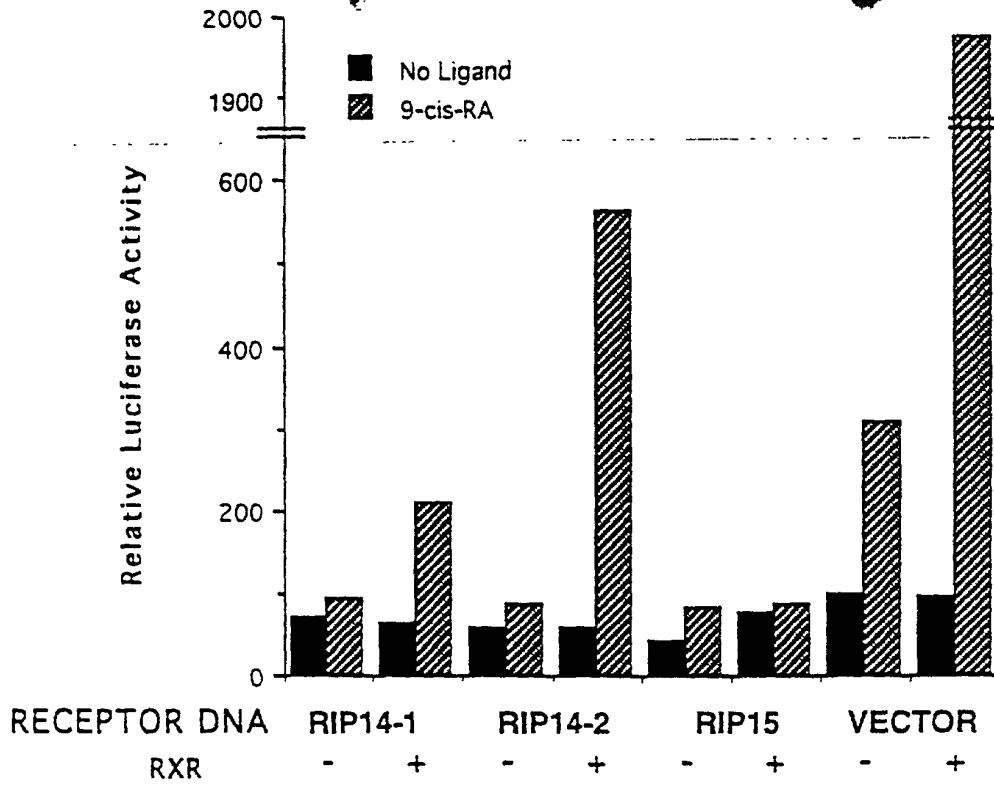
1 2 3 4 5 6 7

RIP14-1	-	+	+	+	+	-	-
RIP14-2	-	-	-	-	-	-	+
RIP14C	-	-	-	-	-	+	-
RXR	-	-	+	+	+	+	+
Competitor	-	-	-	sp	ns	-	-



8 9 10 11 12 13 14

FIG. 9



1

C E P V S G S S G S W D Q S S Q P V F E N -
 GAGAACGTTAAATGTTTTGACAGATGTACTGGCCACTTGGCTGAGCACACACAGTGTGGG
 612 -----+-----+-----+-----+-----+-----+-----+ 671
 CTCTTGCAATTTACAAAAGTGTCTACATGACCGGTGAACCGACTCGTGTGTGTGCACACCC

 C E N V K C F D R C T G H L A E H T Q C G -
 AAGCCACAGGAAAGTACTGGGAGTGGTTATGCTTTTTCCAAAGCTGTCCAGGGTAGGGGG
 672 -----+-----+-----+-----+-----+-----+-----+ 731
 TTCGGTGTCTTTTCATGACCCCTACCAATACGAAAAGGTTTCGACAGGTCCCATCCCCC

 C K P Q E S T G S G Y A F S K A V Q G R G -
 GACACGTCTAGGCAATGCCTTCCTATCCCAGCAGACACAAAAGGTCTCCAGGACACTGGG
 732 -----+-----+-----+-----+-----+-----+-----+ 791
 CTGTGCAGATCCGTTACGGAAGGATAGGGTCGTCTGTGTTTTCCAGAGGTCTGTGACCC

 C D T S R Q C L P I P A D T K G L Q D T G -
 GGCCTGTGCTACTACTGGGGTATTCCATTCTGCCCTGCTGGAGTAGATCCCCAATCAA
 792 -----+-----+-----+-----+-----+-----+-----+ 851
 CCGTGACACGTGATGATGACCCCATAGGTAAGACGGGACGACCTCATCTAGGGTTAGTT

 C G T V H Y Y W G I P F C P A G V D P N Q -
 TACACCAATGTCATTCTCTGCCAGTTAGAGGTTTATCAGAAGAGCCTGAAAATGGCTCAG
 852 -----+-----+-----+-----+-----+-----+-----+ 911
 ATGTGGTTACAGTAAGAGACGGTCAATCTCCAAATAGTCTTCTCGGACTTTTACCGAGTC

 C Y T N V I L C Q L E V Y Q K S L K M A Q -
 AGACAGCTTGTAAAAAAGAGGGTTTGGGGAACCAAGTGTACCTAGACCTCCTTTTCTG
 912 -----+-----+-----+-----+-----+-----+-----+ 971
 TCTGTGCAACAATTTTTTCTCCCAAACCCCTTGGTCACAATGGATCTGGAGGAAAAGAC

 C R Q L V K K R G F G E P V L P R P P F L -
 ATCCAGAATGAATGTGGCCAAGAAGATCAGACTAGTGACAAAAATGAAGGCATCTCAGAA
 972 -----+-----+-----+-----+-----+-----+-----+ 1031
 TAGGTCTTACTTACACCGGTTCTTCTAGTCTGATCACTGTTTTTACTTCCGTAGAGTCTT

 C I Q N E C G Q E D Q T S D K N E G I S E -
 GATATGGGAGATGAAGCCAAAGAGGAAAGGCAGGAATCTAGGGCATCTGTCTGGCACTCA
 1032 -----+-----+-----+-----+-----+-----+-----+ 1091
 CTATACCCCTTACTTCGGTTTCTCCTTTCCGTCCTTAGATCCCGTAGACAGACCGTGAGT

 C D M G D E A K E E R Q E S R A S V W H S -
 GAAACCAAGGATTTTCAAAAAAGTCCAATTAAGCTTGAAACAGAACTTTTGTGGAG
 1092 -----+-----+-----+-----+-----+-----+-----+ 1151
 CTTTGGTTCCTAAAGTTTTTTCAGGTTAATTTTCGAACTTTGTCTTTGAAAACAACCTC

 C E T K D F Q K S P I K S L K Q K L L L E -
 GAAGAACCAACAACCAAGTCGTGGTCAGTCTTCCCAAGGTCTGTTTGTGGAAGAACTCT
 1152 -----+-----+-----+-----+-----+-----+-----+ 1211
 CTTCTTGGTTGTTGGTCAGCACCAGTCAGAAGGGTTCAGACAAACAACCTCTTTGGAGA

 C E E P T T S R G Q S S Q G L F V E E T S -

```

1212 GAAGAAGGTCTGAAGAGTTCGGAAGGAGACAACCTCTGTGCCCACCACGCAAAGCATTGCA
-----+-----+-----+-----+-----+-----+-----+ 1271
CTTCTTCCAGACTTCTCAAGCCTTCCTCTGTTGAGACACGGGTGGTGCCTTTTCGTAACGT

c   E E G L K S S E G D N S V P T T Q S I A -

1272 GCTTTGACCAGTAAGAGAAGTTTAGTTCTTATGCCGAAAGTTCTGCAGAAGAAATCACT
-----+-----+-----+-----+-----+-----+ 1331
CGAAACTGGTCATTCTCTTCAAATCAAGAATACGGCCTTTCAAGACGTCTTCTTTAGTGA

c   A L T S K R S L V L M P E S S A E E I T -

1332 GTTTGCCCTGAGACACAGTTAAGTTTCTTGAACCCCTTGACCTCAATAGAGAAGACTCT
-----+-----+-----+-----+-----+-----+ 1391
CAAACGGGACTCTGTGTCAATTCAAAGGAACCTGGGGAACTGGAGTTATCTCTTCTGAGA

c   V C P E T Q L S F L E P L D L N R E D S -

1392 CCAGATAGCAGAGAGCTCCCCATTGAAGTAAGGATGGCAGTGGGCGATAAGCAGGTTGCT
-----+-----+-----+-----+-----+-----+ 1451
GGTCTATCGTCTCTCGAGGGGTAACCTTCATTCCTACCGTCACCCGCTATTTCGTCCAACGA

c   P D S R E L P I E V R M A V G D K Q V A -

1452 AATAGGGAAGATTGTATGAAGGAAAACCCTCCTCCTGCAGTCTCATCTAGTACCCGGGTA
-----+-----+-----+-----+-----+-----+ 1511
TTATCCCTTCTAACATACTTCCTTTTGGGAGGAGGACGTCAGAGTAGATCATGGGCCCCAT

c   N R E D C M K E N P P P A V S S S T R V -

1512 TCCTGCCCCTGTGTAAACCAAGACTTTCCTCCACAAAGATTGAACAGCATGCCATGTAC
-----+-----+-----+-----+-----+-----+ 1571
AGGACGGGTGACACATTGGTTCTGAAAGGAGGTGTTCTAACTTGTCGTACGGTACATG

c   S C P L C N Q D F P P T K I E Q H A M Y -

1572 TGCAATGGTCTGATGGAGCAGGAAACAGTGTGACTCGGAGACGAAGAGAGGCCAAGAAC
-----+-----+-----+-----+-----+-----+ 1631
ACGTTACCAGACTACCTCGTCCTTTGTCACTGAGCCTCTGCTTCTCTCCGGTTCTTG

c   C N G L M E Q E T V L T R R R R E A K N -

1632 AAGAGTGACGGTCGGACAGCTGCACAGCCGGCTCTGGATGCCAACAGGAAGGAGAAGTGT
-----+-----+-----+-----+-----+-----+ 1691
TTCTCACTGCCAGCCTGTGACGTGTGCGCCGAGACCTACGGTTGTCCTTCCTCTTACA

c   K S D G R T A A Q P A L D A N R K E K C -

1692 TATCTATGTAAGTCCCTGGTTCCACTTGGGGAGTATCAGTGCCATGTGGAGGCCTGTCTC
-----+-----+-----+-----+-----+-----+ 1751
ATAGATACATTACAGGGACCAAGGTGAACCCCTCATAGTCACGGTACACCTCCGGACAGAG

c   Y L C K S L V P L G E Y Q C H V E A C L -

1752 CAGCTTGCAAAGGTTGACAGAGAAGATGGGATTGAAGGGACAAGGAGACCAAGGGTGTGT
-----+-----+-----+-----+-----+-----+ 1811
GTCGAACGTTTCCAACGTCTCTTCTACCCTAACTTCCCTGTTCTCTGGTTCCACACA

c   Q L A K V D R E D G I E G T R R P R V C -

GCACCTGTGGAGGGGAAACAACAGCAGCGGCTGAAGAAGTCAAAGGACAAAGGCCATAGT

```

```

1812 -----+-----+-----+-----+-----+-----+-----+ 1871
      CGTGGACACCTCCCCTTTGTTGTGTCGTCGCCGACTTCTTCAGTTTCCTGTTTCCGGTATCA
c      A P V E G K Q Q Q R L K K S K D K G H S -
      CAAGGCCGACTCCTCAGTCTCTTGGAGCAGTCTGAGCATAGGACCACAGGTGTAGAGAAA
1872 -----+-----+-----+-----+-----+-----+-----+ 1931
      GTTCCGGCTGAGGAGTCAGAGAACCTCGTCAGACTCGTATCCTGGTGTCCACATCTCTTT
c      Q G R L L S L L E Q S E H R T T G V E K -
      AAACCCAAGTATTCGGAAGTAAGAACCTTCAGGATGCCCTCACCAGAGGTGGAAGAGGCT
1932 -----+-----+-----+-----+-----+-----+-----+ 1991
      TTTGGGTTTCATAAGCCTTCATTCTTGGAAGTCTACGGGAGTGGTCTCCACCTTCTCCGA
c      K P K Y S E V R T F R M P S P E V E E A -
      AGCTGCAGCAGAGAGATGCAGAGTACCCTCTCACAGCTCAACTTAAATGAGTCTCCCATC
1992 -----+-----+-----+-----+-----+-----+-----+ 2051
      TCGACGTCGTCTCTCTACGTCTCATGGGAGAGTGTGCGAGTTGAATTTACTCAGAGGGTAG
c      S C S R E M Q S T L S Q L N L N E S P I -
      AAGTCTTTTGTTCCTGTTTCAGAAGCTACAAATTGCTTAGTGGACTTTAAAGAACAGTTT
2052 -----+-----+-----+-----+-----+-----+-----+ 2111
      TTCAGAAAACAAGGACAAAGTCTTCGATGTTTAACGAATCACCTGAAATTTCTTGTCAAA
c      K S F V P V S E A T N C L V D F K E Q F -
      GCTTTCCGGTCACGAACTAAATCAGGCAGGGAAAGGAGGAGAAAATCTTGAATTTCTTGA
2112 -----+-----+-----+-----+-----+-----+-----+ 2171
      CGAAAGGCCAGTGCTTGATTTAGTCCGTCCCTTTCTCTCTTTTAGAACTTAAAGAACT
c      A F R S R T K S G R E R R R K S *
      GACTGGAAGGTTGACCAGAACACACATCGTTGGGTTGATCGTGTTTCAATTAAGTATAGTGG
2172 -----+-----+-----+-----+-----+-----+-----+ 2231
      CTGACCTTCCAACCTGGTCTTGTGTGTAGCAACCAACTAGCACAAGTAATTCATATCACC

      TCTCTAGTTTGTGGTGAGAGTTCTGACCCTGTTGTTATCACCACCAGCACCCATTAGTA
2232 -----+-----+-----+-----+-----+-----+-----+ 2291
      AGAGATCAAACACCACTCTCAAGACTGGGACAACAATAGTGGTGGTCGTGGGTAAGTCAT

      TCCTGGCTTTATATTTTATAAGATCAGTTCAGACAACGTGAATATTATTCTGTTTGAAT
2292 -----+-----+-----+-----+-----+-----+-----+ 2351
      AGGACCGAAATATAAAATATTCTAGTCAAGTCTGTTGACACTTATAATAAGACAACTTA

      TTGCTTATAGTTAAAAATTAAATATATTTATCTTTGTATGAAAAAAAAA
2352 -----+-----+-----+-----+-----+-----+-----+
      AACGAATATCAATTTTAAATTTATATAAATAGAAACATACTTTTTTTTTT

```

GGATGCCCTGGCTGCTCTTGTGGATGCTGCAGCTTCTGCACCCAGATGGATGTTTCCAA
 901 -----+-----+-----+-----+-----+-----+ 960
 CCTACGGGACCGACGAGAACACCTACGACGTGGAAGACGTGGGGTCTACCTACAAAGGTT
 b D A L A A L V D A A A S A P Q M D V S K -
 AACAAAAGAGAGTAAGCATGAAGCTGCCAGGTAGAAAGAAATTTGAGAAGCAGGTCAGC
 961 -----+-----+-----+-----+-----+-----+ 1020
 TTGTTTTCTCTCATTTCGTACTTCGACGGTCCAATCTTCTTTTAACTCTTCGTCCAGTCG
 b T K E S K H E A A R L E E N L R S R S A -
 AGCAGTTAGTGAACAGCAGCAGCTAGAGCAGAAAAACCTGGAGGTGGAGAAGAGATCTGT
 1021 -----+-----+-----+-----+-----+-----+ 1080
 TCGTCAATCACTTGTCTGTCGTCGATCTCGTCTTTTGGACCTCCACCTCTTCTCTAGACA
 b A V S E Q Q Q L E Q K N L E V E K R S V -
 TCAGTGTGTGTGCACTTCTTCAGCCCTTCCAAGTGGCAAGGCCAGCCTCATGCCTCAGT
 1081 -----+-----+-----+-----+-----+-----+ 1140
 AGTCACACACACGTGAAGAAGTCGGAAGGTTACCGTTCGGGGTCGGAGTACGGAGTCA
 b Q C V C T S S A L P S G K A Q P H A S V -
 AGTGTATTCTGAGGCTGGGAAAGATAAAGGGCCTCCTCCAAATCCAGATATGAGGAAGA
 1141 -----+-----+-----+-----+-----+-----+ 1200
 TCACATAAGACTCCGACCTTCTATTTCCCGAGGAGGTTTTAGGTCTATACTCCTTCT
 b V Y S E A G K D K G P P P K S R Y E E E -
 GCTAAGGACCCGAGGGAAGACTACCATTACTGCAGCTAACTTCATAGACGTGACCATCAC
 1201 -----+-----+-----+-----+-----+-----+ 1260
 CGATTCTGGGCTCCCTTCTGATGGTAATGACGTCGATTGAAGTATCTGCACTGGTAGTG
 b L R T R G K T T I T A A N F I D V T I T -
 CCGGCAAATTGCCTCGGACAAGGATGCGAGGGAACGTGGCTCTCAAAGTTCAGACTCTTC
 1261 -----+-----+-----+-----+-----+-----+ 1320
 GGCCGTTTAAACGGAGCCTGTTCTACGCTCCCTTGCACCGAGAGTTTCAAGTCTGAGAAG
 b R Q I A S D K D A R E R G S Q S S D S S -
 TAGTAGCTTGTCTTCTCACAGGTATGAAACGGCTAGTGATGCCATTGAGGTGATAAGTCC
 1321 -----+-----+-----+-----+-----+-----+ 1380
 ATCATCGAACAGAAGAGTGTCATACTTTGCCGATCACTACGGTAACTCCACTATTCAGG
 b S S L S S H R Y E T A S D A I E V I S P -
 CGCCAGCTCACCTGCACCAACCCAGGAAAAGCCACAGGCCTATCAGCCAGACATGGTTAA
 1381 -----+-----+-----+-----+-----+-----+ 1440
 GCGGTTCGAGTGGACGTGGTGGGGTCCTTTTCGGTGTCCGGATAGTCGGTCTGTACCAATT
 b A S S P A P P Q E K P Q A Y Q P D M V K -

1441 GGCAAATCAAGCAGAAAATGAGTCCACTCGACAGTATGAAGGTCCACTGCATCATTATCG
 -----+-----+-----+-----+-----+-----+ 1500
 CCGTTTAGTTCGTCTTTTACTCAGGTGAGCTGTCATACTTCCAGGTGACGTAGTAATAGC
 b A N Q A E N E S T R Q Y E G P L H H Y R -
 GTCCCAGCAGGAATCACCATCTCCACAGCAACAGCCACCACTGCCCCATCTTCCCAGTC
 1501 -----+-----+-----+-----+-----+-----+ 1560
 CAGGGTCGTCCTTAGTGAGGTGTCGTTGTCGGTGGTGACGGGGGTAGAAGGGTCAG
 b S Q Q E S P S P Q Q Q P P L P P S S Q S -
 AGAGGGAATGGGACAGGTGCCCAGGACCCATCGACTGATCACACTTGCTGACCACATCTG
 1561 -----+-----+-----+-----+-----+-----+ 1620
 TCTCCCTTACCCTGTCCACGGGTCTGGGTAGCTGACTAGTGTGAACGACTGGTGTAGAC
 b E G M G Q V P R T H R L I T L A D H I C -
 TCAAATTATCACACAAGATTTTGCTAGAAATCAAGTTCCCTCGCAGCCTTCTACTTCTAC
 1621 -----+-----+-----+-----+-----+-----+ 1680
 AGTTTAATAGTGTGTTCTAAAACGATCTTTAGTTCAAGGGAGCGTCGGAAGATGAAGATG
 b Q I I T Q D F A R N Q V P S Q P S T S T -
 ATTCCAAACTTCACCATCTGCTTTGTCATCCACACCTGTAAGAACTAAAACCTCAAGCCG
 1681 -----+-----+-----+-----+-----+-----+ 1740
 TAAGGTTTGAAGTGGTAGACGAAACAGTAGGTGTGGACATTCTTGATTTTGGAGTTCCGGC
 b F Q T S P S A L S S T P V R T K T S S R -
 CTACAGCCCAGAATCAGTCTCAGACTGTCTTGCATCCCAGACCAGGTCCTAGAGTCTC
 1741 -----+-----+-----+-----+-----+-----+ 1800
 GATGTCGGGTCTTAGTGTGACAGTCTGACAGAACGTAGGGTCTGGTCCAGGATCTCAGAG
 b Y S P E S Q S Q T V L H P R P G P R V S -
 TCCAGAAAATCTTGTGGATAAATCCCGGGGAAGCAGGCCTGGAAAATCTCCAGAGAGGAG
 1801 -----+-----+-----+-----+-----+-----+ 1860
 AGGTCTTTTAGAACACCTATTTAGGGCCCCTTCGTCCGGACCTTTTAGAGGTCTCTCCTC
 b P E N L V D K S R G S R P G K S P E R S -
 TCATATCCCATCAGAGCCCTATGAGCCCATCTCCCCACCCCAAGGCCCTGCTGTGCATGA
 1861 -----+-----+-----+-----+-----+-----+ 1920
 AGTATAGGGTAGTCTCGGGATACTCGGGTAGAGGGGTGGGGTTCCGGGACGACACGTACT
 b H I P S E P Y E P I S P P Q G P A V H E -
 GAAGCAGGACAGCATGTTGCTCTTGTCACAGAGGGGAGTGGACCCTGCTGAGCAAAGGAG
 1921 -----+-----+-----+-----+-----+-----+ 1980
 CTTTCGTCCTGTCGTACAACGAGAACAGTGTCTCCCCTCACCTGGGACGACTCGTTTCCTC
 b K Q D S M L L L S Q R G V D P A E Q R S -
 TGATTCTCGATCACCAGGAAGTATAAGCTACTTGCCTTCATTCTTCACCAAGCTTGAAAG
 1981 -----+-----+-----+-----+-----+-----+ 2040
 ACTAAGAGCTAGTGGTCCTTCATATTCGATGAACGGAAGTAAGAAGTGGTTCGAACCTTC
 b D S R S P G S I S Y L P S F F T K L E S -
 CACATCACCCATGGTTAAATCAAAGAAACAGGAAATTTTTCGTAAGTTGAACTCTTCTGG

2041 -----+-----+-----+-----+-----+-----+ 2100
 GTGTAGTGGGTACCAATTTAGTTTCTTTGTCCTTTAAAAAGCATTCAACTTGAGAAGACC
 b T S P M V K S K K Q E I F R K L N S S G -
 TGGAGGTGACTCTGATATGGCAGCTGCTCAGCCAGGAACAGAGATCTTCAATCTGCCAGC
 2101 -----+-----+-----+-----+-----+ 2160
 ACCTCCACTGAGACTATAACGTCGACGAGTCGGTCTTGTCTCTAGAAGTTAGACGGTTCG
 b G G D S D M A A A Q P G T E I F N L P A -
 AGTTACCACATCAGGTGCAGTGAGCTCAAGAAGCCATTCTTTTGCTGATCCCGCCAGTAA
 2161 -----+-----+-----+-----+-----+ 2220
 TCAATGGTGTAGTCCACGTCAGTCTTCGGTAAGAAAACGACTAGGGCGGTCATT
 b V T T S G A V S S R S H S F A D P A S N -
 CCTTGGTCTAGAAGACATCATCAGAAAGGCTCTCATGGGAAGTTTTGATGATAAAGTTGA
 2221 -----+-----+-----+-----+-----+ 2280
 GGAACCAGATCTTCTGTAGTAGTCTTTCCGAGAGTACCCTTCAAACTACTATTTCAACT
 b L G L E D I I R K A L M G S F D D K V E -
 AGATCATGGTGTGTGTCATGTCCCATCCTGTGGGCATTATGCCTGGTAGTGCCAGCACCTC
 2281 -----+-----+-----+-----+-----+ 2340
 TCTAGTACCACAACAGTACAGGGTAGGACACCCGTAATACGGACCATCACGGTCGTGGAG
 b D H G V V M S H P V G I M P G S A S T S -
 AGTGGTGACGAGCAGCGAGGCACGGAGAGATGAAGGGGAGCCATCACCTCATGCAGGAGT
 2341 -----+-----+-----+-----+-----+ 2400
 TCACCACTGCTCGTCGCTCCGTGCCTCTCTACTTCCCCTCGGTAGTGGAGTACGTCCTCA
 b V V T S S E A R R D E G E P S P H A G V -
 ATGCAAACCAAAGCTGATCAACAAATCAAACAGCAGGAAGTCTAAATCTCCTATTCCTGG
 2401 -----+-----+-----+-----+-----+ 2460
 TACGTTTGGTTCGACTAGTTGTTTAGTTTGTGTCCTTCAGATTTAGAGGATAAGGACC
 b C K P K L I N K S N S R K S K S P I P G -
 GCAAAGCTATTTAGGAAGTGAAGGCCTTCTTCTGTCTCCTCTGTGCATTGAGAAGGTGA
 2461 -----+-----+-----+-----+-----+ 2520
 CGTTTCGATAAATCCTTGACTTTCCGGAAGAAGACAGAGGAGACACGTAAGTCTTCCACT
 b Q S Y L G T E R P S S V S S V H S E G D -
 TTACCACAGGCAGACACCAGGATGGGCATGGGAAGATCGGCCCTCTTCAACAGGTTCTAC
 2521 -----+-----+-----+-----+-----+ 2580
 AATGGTGTCCGTCTGTGGTCCTACCCGTACCCTTCTAGCCGGGAGAAGTTGTCCAAGATG
 b Y H R Q T P G W A W E D R P S S T G S T -
 TCAGTTCCTTACAACCTCTGACCATACGGATGCTCAGCAGTACACCACCTACACAGAT
 2581 -----+-----+-----+-----+-----+ 2640
 AGTCAAGGGAATGTTGGGAGACTGGTATGCCTACGAGTCGTATGTGGTGGATGTGTCTA
 b Q F P Y N P L T I R M L S S T P P T Q I -
 CGCATGCGCCCCATCTGCCATCACCCAAGCAGCTCCACATCAACAGAACCGCATCTGGGA
 2641 -----+-----+-----+-----+-----+ 2700
 CGGTACGCGGGGTAGACGGTAGTGGGTTCTGTCGAGGTGTAGTTGTCTTGGCGTAGACCT

2041
 2101
 2161
 2221
 2281
 2341
 2401
 2461
 2521
 2581
 2641
 2700

b A C A P S A I T Q A A P H Q Q N R I W E -

 GAGGGAGCCTGCCCCGCTCCTCTCAGCGCAGTATGAGACACTGTCTGATAGTGACGACTG

2701 -----+-----+-----+-----+-----+-----+ 2760

 CTCCTTCGGACGGGGCGAGGAGAGTCGCGTCATACTCTGTGACAGACTATCACTGCTGAC

b R E P A P L L S A Q Y E T L S D S D D * -

 AGCTGTGCGTGGGAGAGCGCTCTGGCTTTGGTTTTTATTGAAGATTAAAAAAAAAAAAA

2761 -----+-----+-----+-----+-----+-----+ 2820

 TCGACACGCACCCTCTCGCGAGACCGAAACCAAAAATAACTTCTAAATTTTTTTTTTTTTT

 AA

2821 -- 2822

 TT

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295
1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1379
1380
1381
1382
1383
1384
1385
1386
1387
1388
1389
1390
1391
1392
1393
1394
1395
1396
1397
1398
1399
1400
1401
1402
1403
1404
1405
1406
1407
1408
1409
1410
1411
1412
1413
1414
1415
1416
1417
1418
1419
1420
1421
1422
1423
1424
1425
1426
1427
1428
1429
1430
1431
1432
1433
1434
1435
1436
1437
1438
1439
1440
1441
1442
1443
1444
1445
1446
1447
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457
1458
1459
1460
1461
1462
1463
1464
1465
1466
1467
1468
1469
1470
1471
1472
1473
1474
1475
1476
1477
1478
1479
1480
1481
1482
1483
1484
1485
1486
1487
1488
1489
1490
1491
1492
1493
1494
1495
1496
1497
1498
1499
1500
1501
1502
1503
1504
1505
1506
1507
1508
1509
1510
1511
1512
1513
1514
1515
1516
1517
1518
1519
1520
1521
1522
1523
1524
1525
1526
1527
1528
1529
1530
1531
1532
1533
1534
1535
1536
1537
1538
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1549
1550
1551
1552
1553
1554
1555
1556
1557
1558
1559
1560
1561
1562
1563
1564
1565
1566
1567
1568
1569
1570
1571
1572
1573
1574
1575
1576
1577
1578
1579
1580
1581
1582
1583
1584
1585
1586
1587
1588
1589
1590
1591
1592
1593
1594
1595
1596
1597
1598
1599
1600
1601
1602
1603
1604
1605
1606
1607
1608
1609
1610
1611
1612
1613
1614
1615
1616
1617
1618
1619
1620
1621
1622
1623
1624
1625
1626
1627
1628
1629
1630
1631
1632
1633
1634
1635
1636
1637
1638
1639
1640
1641
1642
1643
1644
1645
1646
1647
1648
1649
1650
1651
1652
1653
1654
1655
1656
1657
1658
1659
1660
1661
1662
1663
1664
1665
1666
1667
1668
1669
1670
1671
1672
1673
1674
1675
1676
1677
1678
1679
1680
1681
1682
1683
1684
1685
1686
1687
1688
1689
1690
1691
1692
1693
1694
1695
1696
1697
1698
1699
1700
1701
1702
1703
1704
1705
1706
1707
1708
1709
1710
1711
1712
1713
1714
1715
1716
1717
1718
1719
1720
1721
1722
1723
1724
1725
1726
1727
1728
1729
1730
1731
1732
1733
1734
1735
1736
1737
1738
1739
1740
1741
1742
1743
1744
1745
1746
1747
1748
1749
1750
1751
1752
1753
1754
1755
1756
1757
1758
1759
1760
1761
1762
1763
1764
1765
1766
1767
1768
1769
1770
1771
1772
1773
1774
1775
1776
1777
1778
1779
1780
1781
1782
1783
1784
1785
1786
1787
1788
1789
1790
1791
1792
1793
1794
1795
1796
1797
1798
1799
1800
1801
1802
1803
1804
1805
1806
1807
1808
1809
1810
1811
1812
1813
1814
1815
1816
1817
1818
1819
1820
1821
1822
1823
1824
1825
1826
1827
1828
1829
1830
1831
1832
1833
1834
1835
1836
1837
1838
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1849
1850
1851
1852
1853
1854
1855
1856
1857
1858
1859
1860
1861
1862
1863
1864
1865
1866
1867
1868
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1880
1881
1882
1883
1884
1885
1886
1887
1888
1889
1890
1891
1892
1893
1894
1895
1896
1897
1898
1899
1900
1901
1902
1903
1904
1905
1906
1907
1908
1909
1910
1911
1912
1913
1914
1915
1916
1917
1918
1919
1920
1921
1922
1923
1924
1925
1926
1927
1928
1929
1930
1931
1932
1933
1934
1935
1936
1937
1938
1939
1940
1941
1942
1943
1944
1945
1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025
2026
2027
2028
2029
2030
2031
2032
2033
2034
2035
2036
2037
2038
2039
2040
2041
2042
2043
2044
2045
2046
2047
2048
2049
2050
2051
2052
2053
2054
2055
2056
2057
2058
2059
2060
2061
2062
2063
2064
2065
2066
2067
2068
2069
2070
2071
2072
2073
2074
2075
2076
2077
2078
2079
2080
2081
2082
2083
2084
2085
2086
2087
2088
2089
2090
2091
2092
2093
2094
2095
2096
2097
2098
2099
2100
2101
2102
2103
2104
2105
2106
2107
2108
2109
2110
2111
2112
2113
2114
2115
2116
2117
2118
2119
2120
2121
2122
2123
2124
2125
2126
2127
2128
2129
2130
2131
2132
2133
2134
2135
2136
2137
2138
2139
2140
2141
2142
2143
2144
2145
2146
2147
2148
2149
2150
2151
2152
2153
2154
2155
2156
2157
2158
2159
2160
2161
2162
2163
2164
2165
2166
2167
2168
2169
2170
2171
2172
2173
2174
217

COMBINED DECLARATION AND POWER OF ATTORNEY

As a below named inventor, I hereby declare that:

My residence, post office address and citizenship are as stated below next to my name.

I believe I am the original, first and sole inventor (if only one name is listed below) or an original, first and joint inventor (if plural names are listed below) of the subject matter which is claimed and for which a patent is sought on the invention entitled RETINOID X RECEPTOR-INTERACTING POLYPEPTIDES AND RELATED MOLECULES AND METHODS the specification of which

☒ is attached hereto.

☐ was filed on _____ as Application Serial No. _____
and was amended on _____.

☐ was described and claimed in PCT International Application No. _____
filed on _____ and as amended under PCT Article 19 on _____.

I hereby state that I have reviewed and understand the contents of the above-identified specification, including the claims, as amended by any amendment referred to above.

I acknowledge the duty to disclose all information I know to be material to patentability in accordance with Title 37, Code of Federal Regulations, §1.56(a).

FOREIGN PRIORITY RIGHTS: I hereby claim foreign priority benefits under Title 35, United States Code, §119 of any foreign application(s) for patent or inventor's certificate or of any PCT international application(s) designating at least one country other than the United States of America listed below and have also identified below any foreign application for patent or inventor's certificate or any PCT international application(s) designating at least one country other than the United States of America filed by me on the same subject matter having a filing date before that of the application(s) of which priority is claimed:

Country	Serial Number	Filing Date	Priority Claimed?
			Yes/No

PROVISIONAL PRIORITY RIGHTS: I hereby claim priority benefits under Title 35, United States Code, §119(e) and §120 of any United States provisional patent application(s) listed below filed by an inventor or inventors on the same subject matter as the present application and having a filing date before that of the application(s) of which priority is claimed:

Serial Number	Filing Date	Status

NON-PROVISIONAL PRIORITY RIGHTS: I hereby claim the benefit under Title 35, United States Code, §120 of any United States application(s) listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in the prior United States application in the manner provided by the first paragraph of Title 35, United States Code, §112, I acknowledge the duty to disclose all information I know to be material to patentability as defined in Title 37, Code of Federal Regulations, §1.56(a) which became available between the filing date of the prior application and the national or PCT international filing date of this application:

Serial Number	Filing Date	Status
08/372,652	January 13, 1995	Pending

I hereby appoint the following attorneys and/or agents to prosecute this application and to transact all business in the Patent and Trademark Office connected therewith: Paul T. Clark, Reg. No. 30,162, Karen L. Elbing, Ph.D. Reg. No. 35,238, Kristina Bieker-Brady, Ph.D. Reg. No. 39,109, Susan M. Michaud, Ph.D. Reg. No. 42,885, Mary Rose Scozzafava, Ph.D., Reg. No.36,268, James D. DeCamp, Ph.D., Reg. No. 43,580.

Address all telephone calls to: Karen L. Elbing, Ph.D. at 617/428-0200.

Address all correspondence to: Karen L. Elbing, Ph.D. at Clark & Elbing LLP, 176 Federal Street, Boston, MA 02110.

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patents issued thereon.

Full Name (First, Middle, Last)	Residence Address (City, State, Country)	Post Office Address (Street, City, State, Country)	Citizenship
David Moore	4523 Birch Street Bellaire, TX 77401	4523 Birch Street Bellaire, TX 77401	United States
Signature:			Date:

Full Name (First, Middle, Last)	Residence Address (City, State, Country)	Post Office Address (Street, City, State, Country)	Citizenship
Wongi Seol			Republic of Korea
Signature:			Date:

Full Name (First, Middle, Last)	Residence Address (City, State, Country)	Post Office Address (Street, City, State, Country)	Citizenship
Hueng-Sik Choi	Hormone Research Center, Chonnam National University, Kwangju, 500-757 KOREA	Hormone Research Center, Chonnam National University, Kwangju, 500-757 KOREA	Republic of Korea
Signature:			Date:

COMBINED DECLARATION AND POWER OF ATTORNEY

As a below named inventor, I hereby declare that:

My residence, post office address and citizenship are as stated below next to my name,

I believe I am the original, first and sole inventor (if only one name is listed below) or an original, first and joint inventor (if plural names are listed below) of the subject matter which is claimed and for which a patent is sought on the invention entitled RETINOID X RECEPTOR-INTERACTING POLYPEPTIDES AND RELATED MOLECULES AND METHODS, the specification of which

■ is attached hereto.

X was filed on JANUARY 13, 1995 as Application Serial No. 08/372,652

and was amended on _____.

□ was described and claimed in PCT International Application No. _____
filed on _____ and as amended under PCT Article 19 on _____.

I hereby state that I have reviewed and understand the contents of the above-identified specification, including the claims, as amended by any amendment referred to above.

I acknowledge the duty to disclose all information I know to be material to patentability in accordance with Title 37, Code of Federal Regulations, §1.56(a).

I hereby appoint the following attorneys and/or agents to prosecute this application and to transact all business in the Patent and Trademark Office connected therewith: Karen F. Lech, Reg. No. 35,238, William E. Booth, Reg. No. 28,933; Karl Bozicevic, Reg. No. 28,807; Barry E. Bretschneider, Reg. No. 28,055; Paul T. Clark, Reg. No. 30,162; Peter J. Devlin, Reg. No. 31,753; Willis M. Ertman, Reg. No. 18,658; David L. Feigenbaum, Reg. No. 30,378; Janis K. Fraser, Reg. No., 34,819; John W. Freeman, Reg. No. 29,066; Timothy A. French, Reg. No. 30,175; Alan H. Gordon, Reg. No. 26,168; Scott C. Harris, Reg. No. 32,030; Mark J. Hebert, Reg. No., 31,766; Gilbert H. Hennessey, Reg. No. 25,759; Charles Hieken, Reg. No. 18,411; Robert E. Hillman, Reg. No. 22,837; G. Roger Lee, Reg. No. 28,963; Steven E. Lipman, Reg. No. 30,011; Gregory A. Madera, Reg. No. 28,878; Ralph A. Mittelberger, Reg. No. 33,195; Ronald E. Myrick, Reg. No. 26,315; Robert C. Nabinger, Reg. No., 33,431; Frank P. Porcelli, Reg. No. 27,374; Eric L. Prael, Reg. No. 32,590; Alan D. Rosenthal, Reg. No. 27,833; Richard M. Sharkansky, Reg. No. 25,800; John M. Skenyon, Reg. No. 27,468; Michael O. Sutton, Reg. No. 26,675; Rene D. Tegtmeyer, Reg. No. 33,567; Hans R. Troesch, Reg. No. 36,950; John N. Williams, Reg. No. 18,948; Gary A. Walpert, Reg. No. 26,098; Dorothy P. Whelan, Reg. No., 33,814; and Charles C. Winchester, Reg. No. 21,040.

Address all telephone calls to Karen F. Lech at telephone number 617/542-5070.

Address all correspondence to Paul T. Clark, Fish & Richardson P.C., 225 Franklin Street, Boston, MA 02110-2804.

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patents issued thereon.

Full Name of Inventor: DAVID MOORE

Inventor's Signature:  Date: 3/27/95

Residence Address: One Isaac Sprague Dr., Hingham, MA 02043

Citizen of: UNITED STATES OF AMERICA

Post Office Address: ONE ISAAC SPRAGUE DRIVE, HINGHAM, MA 02043

EXTRA COPY

Revised: August 24, 1994 (391DECL.MRG)

COMBINED DECLARATION AND POWER OF ATTORNEY CONTINUED

Full Name of Inventor: WONGI SEOL

Inventor's Signature: Wongi Seol Date: 3/27/95

Residence Address: 812 Memorial Dr. A1105 Cambridge, MA 02139

Citizen of: Republic of Korea

Post Office Address: Same

Full Name of Inventor: HUENG-SIK CHOI

Inventor's Signature: Hueng-Sik Choi Date: 4/1/95

Residence Address: Lucky HanHa Rm Apt 101-Dong, 302-Ho, Yusong Gu, Taejeon

Citizen of: Republic of KOREA KOREA (South)

Post Office Address: Immunology Division, Molecular Biomedicine Group
KORZA Research Institute of Bioscience & Biotechnology